

ICCGI 2019

The Fourteenth International Multi-Conference on Computing in the Global Information Technology

ISBN: 978-1-61208-718-4

June 30 – July 4, 2019

Rome, Italy

ICCGI 2019 Editors

Claus-Peter Rückemann, Westfälische Wilhelms-Univ. Muenster (WWU) / DIMF / Leibniz Universitaet Hannover, Germany

Kazuhito Sato, Akita Prefectural University, Japan

ICCGI 2019

Foreword

The Fourteenth International Multi-Conference on Computing in the Global Information Technology (ICCGI 2019), held between June 30 – July 4, 2019 - Rome, Italy, continued a series of international events covering a large spectrum of topics related to global knowledge technologies, computation, mechanisms, cognitive patterns, concerning thinking, communications, user-centric approaches, nanotechnologies, and advanced networking and systems. The conference topics focus on challenging aspects in the next generation of information technology and communications related to the computing paradigms (mobile computing, database computing, GRID computing, multi-agent computing, autonomic communication computing, evolutionary computation) and and networking and telecommunications technologies (mobility, networking, bio-technologies, autonomous systems, image processing, Internet and web technologies), towards secure, self-defendable, autonomous, privacy-safe, and context-aware scalable systems.

This conference intended to expose the scientists to the latest developments covering a variety of complementary topics, aiming to enhance one's understanding of the overall picture of computing in the global information technology.

The integration and adoption of IPv6, also known as the Next Generation of the Internet Protocol, is happening throughout the World at this very moment. To maintain global competitiveness, governments are mandating, encouraging or actively supporting the adoption of IPv6 to prepare their respective economies for the future communication infrastructures. Business organizations are increasingly mindful of the IPv4 address space depletion and see within IPv6 a way to solve pressing technical problems while IPv6 technology continues to evolve beyond IPv4 capabilities. Communications equipment manufacturers and applications developers are actively integrating IPv6 in their products based on market demands.

IPv6 continues to represent a fertile area of technology innovation and investigation. IPv6 is opening the way to new successful research projects. Leading edge Internet Service Providers are guiding the way to a new kind of Internet where any-to-any reachability is not a vivid dream but a notion of reality in production IPv6 networks that have been commercially deployed. National Research and Educational Networks together with internationally known hardware vendors, Service Providers and commercial enterprises have generated a great amount of expertise in designing, deploying and operating IPv6 networks and services. This knowledge can be leveraged to accelerate the deployment of the protocol worldwide.

We take here the opportunity to warmly thank all the members of the ICCGI 2019 Technical Program Committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to ICCGI 2019. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the ICCGI 2019 organizing

committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that ICCGI 2019 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the area of computing in the global information technology.

We are convinced that the participants found the event useful and communications very open. We also hope that Rome provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city.

ICCGI 2019 Chairs

ICCGI Steering Committee

Carlos Becker Westphall, Universidade Federal de Santa Catarina, Brazil Constantin Paleologu, University Politehnica of Bucharest, Romania Elena Ravve, Ort-Braude College - Karmiel, Israel Ulrich Reimer, University of Applied Sciences St. Gallen, Switzerland Juho Mäkiö, Hochschule Emden / Leer, Germany Abdel Lisser, Université Paris Sud, France Robert Bestak, Czech Technical University in Prague, Czech Republic

ICCGI Industry/Research Advisory Committee

John Terzakis, Intel, USA Isabel Seruca, Portucalense University, Portugal Yasushi Kambayashi, Nippon Institute of Technology, Japan Arno Leist, Institute of Natural and Mathematical Sciences | Massey University Auckland, New Zealand Peter Kieseberg, St. Poelten University of Applied Sciences, Austria Isabel Muench, BSI - German Federal Office for Information Security, Germany Zeeshan Pervez, University of the West of Scotland, UK

ICCGI 2019

Committee

ICCGI Steering Committee

Carlos Becker Westphall, Universidade Federal de Santa Catarina, Brazil Constantin Paleologu, University Politehnica of Bucharest, Romania Elena Ravve, Ort-Braude College - Karmiel, Israel Ulrich Reimer, University of Applied Sciences St. Gallen, Switzerland Juho Mäkiö, Hochschule Emden / Leer, Germany Abdel Lisser, Université Paris Sud, France Robert Bestak, Czech Technical University in Prague, Czech Republic

ICCGI Industry/Research Advisory Committee

John Terzakis, Intel, USA Isabel Seruca, Portucalense University, Portugal Yasushi Kambayashi, Nippon Institute of Technology, Japan Arno Leist, Institute of Natural and Mathematical Sciences | Massey University Auckland, New Zealand Peter Kieseberg, St. Poelten University of Applied Sciences, Austria Isabel Muench, BSI - German Federal Office for Information Security, Germany Zeeshan Pervez, University of the West of Scotland, UK

ICCGI 2019 Technical Program Committee

António Abreu, Polytechnic Institute of Lisbon, Portugal Nadine Akkari, King Abdulaziz University, Jeddah, Saudi Arabia Konstantin Aksyonov, Ural Federal University, Russia Cristina Alcaraz, University of Malaga, Spain Ala Alluhaidan, Claremont Graduate University, USA Diana Andone, Politehnica University of Timisoara, Romania Antonova Anna, Ural Federal University, Russia Ruzaini Abdullah Arshah, University Malaysia Pahang, Malaysia Zahra Asebriy, Cadi Ayyad University, Morocco Jocelyn Aubert, Luxembourg Institute of Science and Technology (LIST), Luxembourg Sara Balderas-Díaz, University of Granada, Spain Michaela Baumann, University of Bayreuth, Germany Carlos Becker Westphall, Universidade Federal de Santa Catarina, Brazil Hatem Ben Sta, University of Tunis at El Manar, Tunisia Robert Bestak, Czech Technical University in Prague, Czech Republic Ateet Bhalla, Independent Consultant, India Dorota Bielinska-Waz, Medical University of Gdansk, Poland Fernando Bobillo, University of Zaragoza, Spain

Mihai Boicu, George Mason University, USA Eugen Borcoci, University "Politehnica" of Bucharest (UPB), Romania Leonardo Botega, University Centre Eurípides of Marília (UNIVEM), Brazil Jean-Louis Boulanger, CERTIFER, France Christos J. Bouras, University of Patras, Greece Rustam Burnashev, Kazan Federal University, Russia Xiaoqiang Cai, The Chinese University of Hong Kong (Shenzhen), P. R. China Maiga Chang, Athabasca University, Canada Mu-Song Chen, Da-Yeh University, Taiwan Albert M. K. Cheng, University of Houston, USA Yung Ryn (Elisha) Choe, Sandia National Laboratories, Livermore, USA Rebeca Cortazar, University of Deusto, Spain Beata Czarnacka-Chrobot, Warsaw School of Economics, Poland Mirela Danubianu, "Stefan cel Mare" University of Suceava, Romania Moussa Diallo, University Cheikh Anta Diop of Dakar, Senegal António Dourado, University of Coimbra, Portugal Hans-Dieter Ehrich, Technische Universität Braunschweig, Germany Samuel Ekundayo, Eastern Institute of Technology, New Zealand Mohammed Nabil El Korso, Paris Nanterre University, France Ahmed El Oualkadi, Abdelmalek Essaadi University, Morocco Sylvia Encheva, Western Norway University of Applied Sciences, Norway Manuel Filipe Santos, University of Minho, Portugal Rita Francese, Università di Salerno - Fisciano, Italy Dirk Frosch-Wilke, University of Applied Sciences Kiel, Germany Félix J. García, University of Murcia, Spain Joy Garfield, University of Worcester, UK Olga Georgieva, Sofia University "St. Kl. Ohridski", Bulgaria Katja Gilly de la Sierra - Llamazares, Universidad Miguel Hernández, Spain Apostolos Gkamas, University Ecclesiastical Academy of Vella of Ioannina, Greece Mikhail Gofman, California State University, Fullerton Mario Goldenbaum, Princeton University, USA William Grosky, University of Michigan-Dearborn, USA Carlos Guerrero, University of Balearic Islands, Spain Gabriel Guerrero-Contreras, University of Cádiz, Spain Christophe Guyeux, Femto ST Institute | UMR 6173 CNRS | University of Bourgogne Franche-Comté, France Maki K. Habib, The American University in Cairo, Egypt Petr Hanáček, Brno University of Technology, Czech Republic Fei Hao, Shaanxi Normal University, China Abolfazl Hashemi, University of Texas at Austin, USA Wladyslaw Homenda, Warsaw University of Technology, Poland Tzung-Pei Hong, National University of Kaohsiung, Taiwan Wei-Chiang Hong, School of Education Intelligent Technology - Jiangsu Normal University, China Vitalyi Igorevich Talanin, Institute of Economics & Information Technologies, Zaporozhye, Ukraine Kyoko Iwasawa, Takushoku University, Tokyo, Japan Motoi Iwashita, Chiba Institute of Technology, Japan Abdelilah Jilbab, Mohammed V University in Rabat, Morocco Maria João Ferreira, Universidade Portucalense, Portugal

Imed Kacem, Université de Lorraine, France Yasushi Kambayashi, Nippon Institute of Technology, Japan Georgios Kambourakis, University of the Aegean, Greece Jean Robert Kala Kamdjoug, Université Catholique d'Afrique Centrale | Institut Catholique de Yaoundé, Cameroon Dimitris Kanellopoulos, University of Patras, Greece Paweł Karczmarek, Lublin University of Technology, Poland Andrzej Kasprzak, Wroclaw University of Technology, Poland Peter Kieseberg, St. Poelten University of Applied Sciences, Austria Wojciech Kmiecik, Wroclaw University of Technology, Poland Leszek Koszalka, Wroclaw University of Science and Technology, Poland Piotr A. Kowalski, AGH University of Science and Technology / Systems Research Institute - Polish Academy of Sciences, Poland Panos Kudumakis, Queen Mary University of London, UK Bobby Law, Glasgow Caledonian University, Scotland Arno Leist, Institute of Natural and Mathematical Sciences | Massey University Auckland, New Zealand Yin Leng Tan, Henley Business School | University of Reading, UK Isaac Lera, Universitat de les Illes Balears, Spain Quan-Lin Li, Yanshan University, China Abdel Lisser, Université Paris Sud, France Fernando Luís Almeida, University of Porto & INESC TEC, Portugal Ivan Luković, University of Novi Sad, Serbia Mary Luz Mouronte López, Universidad Francisco de Vitoria - Madrid, Spain Stephane Maag, Institut Mines Telecom / Telecom SudParis, France José Machado, University of Minho, Portugal Olaf Manuel Maennel, TalTech University, Estonia Paulo Maio, School of Engineering (ISEP) of Polytechnic of Porto (IPP), Portugal Juho Mäkiö, Hochschule Emden / Leer, Germany Pedro Malagón, Universidad Politécnica de Madrid, Spain Giuseppe Mangioni, University of Catania, Italy Goreti Marreiros, Instituto Superior de Engenharia do Porto / GECAD, Portugal Constantino Martins, School of Engineering of the Polytechnic Institute of Porto (ISEP / IPP), Portugal Natarajan Meghanathan, Jackson State University, USA Nuno Melão, Instituto Politécnico de Viseu - Escola Superior de Tecnologia e Gestão de Viseu, Portugal Abdelkrim Meziane, Research Center in Scientific and Technical information CERIST, Algeria Angelos Michalas, TEI of Western Macedonia, Kastoria, Greece Martin Misut, University of Economics in Bratislava, Slovak Republic Fernando Moreira, Universidade Portucalense Infante D. Henrique, Portugal Mawloud Mosbah, University 20 Août 1955 of Skikda, Algeria Paulo Moura Oliveira, UTAD University | INESC-TEC, Portugal Isabel Muench, BSI - German Federal Office for Information Security, Germany Gyu Myoung Lee, Liverpool John Moores University, UK Antonio Navarro, Universidad Complutense de Madrid, Spain Rafael Nogueras, University of Malaga, Spain Isabel Novo-Corti, Universidade da Coruña, Spain Joshua C. Nwokeji, Gannon University - Erie Pennsylvania, USA Constantin Paleologu, University Politehnica of Bucharest, Romania Thanasis Papaioannou, Athens University of Economics and Business (AUEB), Greece

Isabel Pedrosa, Coimbra Business School - ISCAC, Portugal Manuel Pérez Cota, Universidad de Vigo, Spain Bernhard Peischl, Institut for Softwaretechnologie | Technische Universität Graz, Austria José Henrique Pereira São Mamede, Universidade Aberta, Portugal Zeeshan Pervez, University of the West of Scotland, UK Iwona Pozniak-Koszalka, Wroclaw University of Science and Technology, Poland Shaojie Qiao, College of Information Security Engineering - Chengdu University of Information Technology, China Maria Paula Queluz, Instituto de Telecomunicações | Instituto Superior Técnico/TULisbon, Portugal Kornelije Rabuzinm, University of Zagreb, Croatia Rahim Rahmani, University of Stockholm, Sweden Elena Ravve, Ort-Braude College - Karmiel, Israel Marek Reformat, University of Alberta, Canada Ulrich Reimer, University of Applied Sciences St. Gallen, Switzerland Fernando Reinaldo Ribeiro, Polytechnic Institute of Castelo Branco, Portugal Luis Paulo Reis, University of Minho, Portugal Éric Renault, Institut Mines-Télécom | Télécom SudParis, France Feliz Ribeiro Gouveia, Universidade Fernando, Portugal Ruben Ricart-Sanchez, University of the West of Scotland, UK Michele Risi, University of Salerno, Italy Alvaro Rocha, University of Coimbra, Portugal Michele Ruta, Technical University of Bari, Italy Shrirang Sahasrabudhe, University North Carolina Greensboro, USA Ozgur Koray Sahingoz, Turkish Air Force Academy, Turkey Peter Schartner, Alpen-Adria-Universität Klagenfurt, Austria Ina Schiering, Institut für Information Engineering | Ostfalia Hochschule für angewandte Wissenschaften, Germany Wieland Schwinger, Johannes Kepler University Linz (JKU), Austria Floriano Scioscia, Polytechnic University of Bari, Italy Florence Sèdes, Université Toulouse 3, France Isabel Seruca, Portucalense University, Portugal Martina Šestak, University of Zagreb, Croatia Marc Sevaux, Université de Bretagne-Sud, France Ashok Sharma, TechMahindra, India Luis Silva Rodrigues, Politécnico do Porto / ISCAP, Portugal Kathryn E. Stecke, University of Texas at Dallas, USA Yeong-Tae Song, Towson University, USA Pedro Sousa, University of Minho, Portugal Ali Taalimi, The University of Tennessee at Knoxville, USA John Terzakis, Intel, USA Chrisa Tsinaraki, EU JRC, Italy Ion Tutanescu, University of Pitesti, Romania Laura Varela-Candamio, University of A Coruna, Spain Radu Adrian Vasiu, Politehnica University of Timisoara, Romania Michael Vassilakopoulos, University of Thessaly, Greece Paula Viana, Polytechnic of Porto, School of Engineering and INESC TEC, Portugal Piotr Waz, Medical University of Gdansk, Poland Ewa Więcek-Janka, Poznan University of Technology, Poland

Yair Wiseman, Bar-Ilan University, Israel Ouri Wolfson, University of Illinois, USA Kuan Yew Wong, Universiti Teknologi Malaysia (UTM), Malaysia Mudasser F. Wyne, National University, USA Mohammad Yahya H. Al-Shamri, Ibb University, Yemen Aleš Zamuda, University of Maribor, Slovenia

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

| Analysis of Cognitive Functions in the Encountered State of Driving Topics on Expressways Kazuhito Sato, Yuki Oomomo, Hirokazu Madokoro, Momoyo Ito, and Sakura Kadowaki | 1 |
|--|----|
| Survey on Smart Web Crawler Algorithm Based on Semantic Search Engine Muneer Bani Yassein, Ismail Hmeidi, Yaser Khamayseh, Wail Mardini, Heba Al-Jarrah, and Dragana Krstic | 9 |
| Innovation as a Structural Attribute of Company Competitiveness Oswaldo Luiz Agostinho | 13 |

Analysis of Cognitive Functions in the Encountered State of Driving Topics on Expressways

Kazuhito Sato, Yuki Oomomo, Hirokazu Madokoro Department of Machine Intelligence and Systems Engineering, Faculty of Systems Science and Technology, Akita Prefectural University Yurihonjo, Japan Email: ksato, m18a016, madokoro @akita-pu.ac.jp

Momoyo Ito

Institute of Technology and Science, Tokushima University Tokushima, Japan Email: momoito@is.tokushima-u.ac.jp

Abstract - For this study, we defined four situations that are likely to trigger accidents on highways as driving topics, i.e., breakdown vehicles, sudden appearance of small animals, falling objects, and lane decreases. Carrying out running experiments controlling the time period (daytime or nighttime) and the traffic flow (overtaking vehicle: Yes or No), we studied the comparative analysis for each driving topic, focusing on driving characteristics of young drivers and elderly drivers concerning to "cognitions", "judgments", and "operations". Using evaluation experiments, we try to analyze the driving characteristics of elderly drivers for driving topics encountered on highways based on comparing to young drivers.

Keywords - driving behavior; human engineering; elderly drivers; driving topic; highway

I. INTRODUCTION

Although the number of deaths due to traffic accidents in Japan is on a downward trend, traffic accidents caused by erroneous operations by elderly drivers and reverse runs of highways have become particularly problematic in recent years. Regarding deaths by age group due to traffic accidents in June, 2015, elderly people over 65 years old (1.006 people, the composition ratio was 53.1%) were the most frequent, followed by the 50s (200 people, the composition ratio was 10.6%), the 40s (169 people, the composition ratio was 8.9%) in this order [1]. Looking at the trend over the past 10 years, although the overall trend is in a decreasing trend (0.60 times in 2005), the rate of decrease in elderly people is low, it is reported that elderly people are the most frequent for eight consecutive years [1], even in the number of fatal accidents by age group such as automobile drivers. As described above, traffic accidents caused by elderly drivers are relatively frequent and have a high risk of leading to serious injury, so it is urgent to take countermeasures towards the arrival of a super aging society coming close to us.

Sakura Kadowaki Smart Design Corp. Akita, Japan Email: sakura@smart-d.jp

As a factor of traffic accidents by elderly drivers, there may be common points unique to the elderly, such as slowing down a series of driving operations related to cognitions, judgments and operations, miss operations of handles and brakes due to distribution of attention to multiple tasks and concentration, momentary driving due to carelessness and distraction. Recently, various traffic accident preventive safety systems have been studied, as symbolized by automatic driving efforts, and practical application of these systems has been accelerating. However, most of those systems collect and analyze external information of the vehicle, and there is a problem, which cannot adaptively deal with the driving characteristics of individual drivers. Also, with preventive safety technology based on a standard general driver, it is difficult to deal with the slowing down of the driving behavior peculiar to the elderly found in recognition, judgment and operation. Therefore, approaches to preventive safety systems dedicated to driving characteristics of elderly drivers are indispensable.

In this research, we define the four driving situations that are likely to trigger accidents as driving topics, such as breakdown vehicles with road shoulder, sudden appearance of small animals, falling objects, and lane decrease, especially for highways with high mortality rate and easy to lead to serious accidents. Then, we carry out running experiments controlling the time period (daytime or nighttime) and the traffic flow (overtaking vehicle: Yes or No). Specifically, paying attention to a series of driving behaviors related to cognitions, judgments and operations for each driving topic, and we will compare and analyze the driving characteristics of young drivers and elderly drivers. Acquiring the head posture, face orientation, eye-gaze, which are less burdens on the driver as body information indicating the driving characteristics by unconstrained means, heartbeat as biological information, handle, brake, and accelerator as each operation, we will clarify the driving

characteristics of elderly drivers for driving topics encountered on highways.

This paper is presented as follows. We review related work to clarify the position of this study in Section II. Section III presents a definition of the experimental protocols, driving topics, and running scenarios. In Section IV, we examine the correspondence relationship between measurement points of driver reaction and eye-gaze information, focusing on the reaction time required for each action, i.e., cognitions, judgments, and operations for the driving topics. Additionally, we compare the average of reaction times of whole drivers, elderly drivers, and young drivers at each measurement point for "falling objects" which is one of the driving topics. Finally, we present conclusions and intentions for future work in Section V.

II. RELATED WORKS

The cause of fatal accidents of elderly drivers is mostly unsuitable for driving operation and is attributed to physical characteristics, psychological characteristics, driving and social characteristics. characteristics, Physical characteristics indicate declining physical function such as vision and exercise ability. As such assistive technologies, systems reducing the operation load of steering, systems automatically reducing the speed and damage when there is a possibility of collision, they have been putting into practical use [2]. The psychological characteristics mean that elderly drivers are not good at parallel processing of multiple information and have a tendency to become selforiented. The driving characteristics refer to mismatch of consciousness and behaviors caused by "accustomed" or "driving that would be" based on past experiences. Additionally, the social characteristics refer to differences in characteristics by generations, such as a decline in communication skills and the influence of motorization [3]. As solutions to these characteristics of elderly drivers, the followings have been reported. It is possible to reform consciousness and motivate for safe driving by having each elderly participate in efforts for traffic safety with their own will rather than passive [3]. The elderly drivers are able to improve attention by pointing and designating the signs to recognize.

Furthermore, as a study on the cognitive function of elderly drivers, Takahara et al. [5] analyzed the characteristics at the temporary stop location, and demonstrated the effectiveness of their system through the development of a voice guidance type temporary stop support system. In addition, Iida et al. [6] constructed the hypothesis of the reverse running process as an example of the elderly driver's accidents on the expressway, they confirmed the psychological state and the road composition of the elderly driver when reverse running is easy to occur. However, these researches have not been analyzed focusing on the situations (driving topics) which are easy to induce accidents on highways, such as breakdown car on road

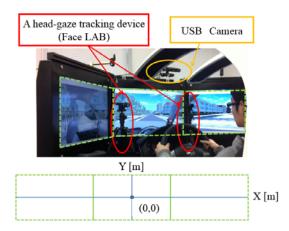


Figure 1. Experimental system for measuring driver behaviors.

shoulder, running-out of small animals, falling objects, and reduced lane during work. In this study, through driving experiments controlled the time zone and traffic flow, we analyze the series of driving behaviors related to cognitions, judgments and operations of these driving topics, and clarify the driving characteristics of elderly drivers.

III. METHODS

In this section, we define the experimental protocols and four driving topics easy to occur accidents. Next, we present running scenarios controlling daytime or nighttime, the presence or absence of an overtaking vehicle.

A. Experimental Systems

Though many people do driving behaviors every day, there are many difficulties to clarify individual driving characteristics from actual behaviors on the road in real environments. For example, driving behaviors vary depending on the road environment and traffic conditions at that time, and their conditions on the actual road cannot be reproduced constantly. Therefore, even if the behavior varies, it is not possible to clearly distinguish whether it is a variation due to a difference in traffic conditions or a variation among individuals.

This study used a Driving Simulator (DS) to assess driving behaviors for freely set road environments and traffic conditions that affect driver behaviors. Figure 1 portrays the experimental system configuration used to measure driver behaviors. The DS used for experiments has platforms corresponding to compact and six-axis motion, which is equipped with ordinary cars. The DS has three color liquid crystal displays mounted in the front of the cabin, and has a function reproducing pseudo-driving environments that are freely configurable to horizontal viewing angles. As Figure 1 shows, to measure body information, such as head poses, face orientations, and evegaze movements without restraining drivers, we installed cameras to the left and right in the center of the three-color liquid crystal monitors mounted in front of the cabin. Additionally, we set an infrared pod on top of the instruments in front of the cabin. Here the camera heads and infrared pod are input-based sensors of a head-gaze tracking device (FaceLAB; Ekstreme Machine Corp.). Incidentally, through preliminary test runs by multiple subjects, we confirmed that installation of the stereo camera head and infrared pod does not interfere with visibility during driving operations. Furthermore, in order to capture the facial expressions of drivers, an USB camera (Xtion pro Live) was installed at the top of the liquid crystal monitor on the center in front of the cabin.

B. Experimental Protocols

Figure 2 depicts the experimental protocol outline. Initially, as individual characteristics of each subject, we conducted an examination of the following questionnaire methods: attitude, oriented, and concept to work on driving were performed using the driving style check sheet [20]. Regarding the types of operation burdens that were strongly felt, they were performed using a driving load sensitivity check sheet [21]. In one running test, for each target subject wearing a heart rate monitor (RS800CX; Polar), we measured the instantaneous heart rate of a normal state during 1 min in advance. Next, to improve the measurement accuracy for face orientations and eye-gaze movements of each participant, we calibrated the cameras of the head-gaze tracking device (FaceLAB). We recorded a face video while driving with the USB camera (Xtion Pro Live; ASUS Corp.) to analyze the facial expressions of the subject. After these preparations, each participant ran along the four running scenarios described later in Section III.C, by synchronizing the time base of each measuring device. Finally, using a questionnaire that specifically examines driving topics

occurring at random, subjective reviews, a four-stage check, were also conducted when a driving topic occurred. After obtaining the approval of Akita Prefectural University Research Ethics Board, the experiment contents for all subjects were explained fully to participants in advance. We obtained written consent of participants. From each, we also obtained an agreement to publish a face image along with the consent to experimental participation. Subjects were 10 young drivers (A, B, C, D, E, F, G, H, I, Q: average age 22 years) and nine elderly drivers (J, K, L, M, N, O, P, R, S: average age 62 years) in total 19 people.

C. Driving topics and Running scenarios

The driving course used in the experiment is a straight course with no lane changes of two lanes on one side

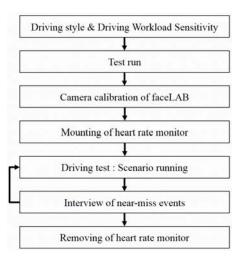


Figure 2. Outline of experimental protocols.

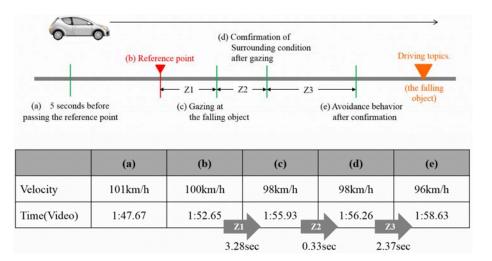
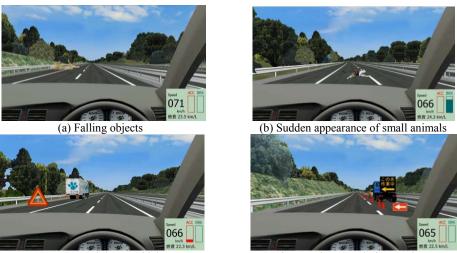


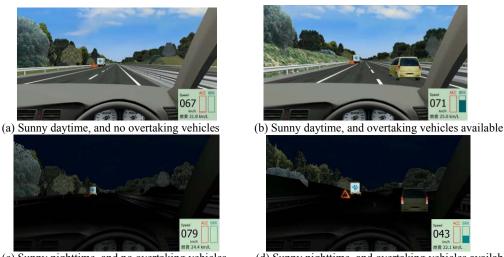
Figure 3. Measurement points of driver reaction to driving topics.



(c) Breakdown vehicles

(d) Lane decrease during work

Figure 4. Driving topics for running test.



(c) Sunny nighttime, and no overtaking vehicles

(d) Sunny nighttime, and overtaking vehicles available

Figure 5. Running scenarios.

simulating the Tohoku Expressway, and starts running from the driving scene that converges from the acceleration lane to the expressway. Driving topics occurred in one running were set randomly, and the measurement points of the driver's reaction for each driving topic were set as shown in Figure 3. Figure 4 presents the state of each driving topic. The falling object in Figure 4 (a) is installed in the center of the left lane, and interferes with the running course. The sudden appearance of small animals shown in Figure 4 (b) was set so that when the vehicle passed the fixed point, it emerged from the front left side of the vehicle and traversed to the right so as to close the driving path of the concerned vehicle. The breakdown vehicle in Figure 4 (c) simulated a state where a large truck was parked on the shoulder due to a failure. With the lane decrease in Figure 4 (d), the overtaking lane is unavailable due to in-work and decreases from 2 lanes to 1 lane. In the case where each driving topic and the concerned vehicle come into contact, characters of "collision" are displayed on the front screen, but running experiments can be continued.

Subsequently, we present an overview of running scenarios in the following sentences. To take into account the differences in visibility due to time periods, we set the two conditions of daytime and nighttime, as well as the presence or absence of an overtaking vehicle as traveling conditions when encountering each driving topic. Figure 5

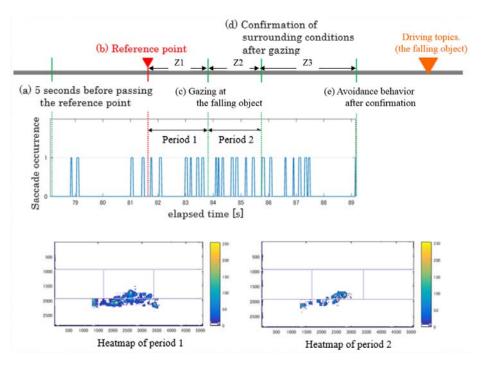


Figure 6. Time series changes of heatmaps and saccades corresponding to each measurement point.

presents four types of driving scenarios that controlled time periods (i.e., daytime or nighttime) and traffic flows (i.e., overtaking vehicles: Yes or No). The order of occurrence of each driving topic was two patterns of type A (i.e., falling objects \rightarrow sudden appearance of small animals \rightarrow breakdown vehicles \rightarrow lane decrease) and type B (i.e., breakdown vehicles \rightarrow sudden appearance of small animals \rightarrow falling objects \rightarrow lane decrease). The running time for one scenario is about 7 to 10 minutes, and we gave instructions to each subject to observe the speed limit of 100 (km / h). Additionally, maximum speed was automatically limited to 120 (km / h) on DS side.

IV. EXPERIMENTAL RESULTS AND ARGUMENTS

Based on safety confirmation behaviors due to the driving topics, we analyzed reaction times of whole drivers, elderly drivers, and young drivers at each measurement point. Finally, we discussed the differences of driving characteristics between elderly drivers and young drivers, focusing on "cognitions", "judgments", and "operations".

A. Measurement points and driver response for driving topics

Focusing attention on eye-gaze behaviors for driving topics and safety confirmation behaviors due to lane change, we will examine the correspondence relationship between measurement points of driver reaction and eye-gaze information (i.e., heat maps, saccades) as shown in Figure 3. Figure 6 presents the time series changes of heat maps and saccades corresponding to each measurement point for the falling object. Each heat map represents the degree of gaze concentration in each of the following sections. They are from b) passing the reference point to c) falling object gazing, from c) falling object gazing to d) checking surrounding conditions, from d) checking surrounding conditions to e) avoidance actions. In addition, the time series changes of saccades are divided into two sections, i.e., from a) 5 seconds before passing the reference point to b) passing the reference point, from b) passing from the reference point to e) avoidance action. Focusing on the time series changes of heat maps and saccades in Figure 6, we are able to confirm driver states related to the attention to the falling object and the surrounding situation confirmation due to lane change. Consequently, based on the vehicle speed of b) the reference point during the passage, it is possible to estimate the reaction time required for each action, i.e., cognitions, judgments, and operations for the falling object.

B. Analysis of Driver Response to Falling Objects

Figure 7 shows the average of reaction times of whole drivers, elderly drivers, and young drivers at each measurement point for "falling objects" which is one of the driving topics. In Figure 7, the respective reaction times are as follows.

1) from b) passing the reference point to c) falling object gazing,

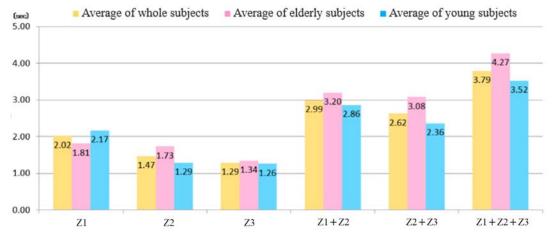


Figure 7. Driver's reaction time to falling objects.

- 2) from c) falling object gazing to d) surrounding situation confirmation,
- 3) from d) surrounding situation confirmation to e) avoidance behavior

When analyzing a series of driving behaviors related to cognitions, judgments and operations for falling objects, the elderly drivers generally tended to have longer reaction times than those of young drivers. Furthermore, by classifying each driver based on the average of the whole subjects and paying attention to the relationship with the driving style, we analyzed the existence of the following four groups.

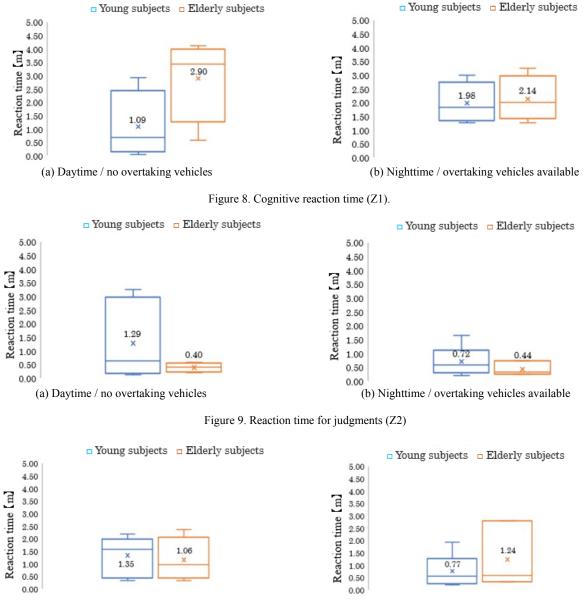
- Group A: Discovery and situation confirmation are quick, allowance for operation (i.e., cognition, judgment, and operation are all fast)
- Group B: Discovery is late, but situation confirmation and handling are quick (i.e., cognition is slow, but judgment and operation are fast)
- Group C: discovery and situation confirmation is late, operation is gradual (i.e., cognition, judgment, and operation are gentle)
- · Group D: all reactions are average (i.e., cognition, judgment, and operation are all average)

The young drivers are classified into Group A and Group D, whereas the elderly drivers are more likely to be classified as Group B and Group C. In addition, we believe that the characteristics of driving behaviors will shift from group B to group C as the aging of driver progresses.

C. Comparison of elderly drivers and young drivers

In Section IV.B, as the result of analyzing a series of reaction times related to cognitions, judgments and operations for falling objects, we confirmed that each driver has a strong tendency to be classified as one of four groups (i.e., A, B, C, and D). In order to further clarify the driving characteristics of elderly drivers, we try to analyze each

reaction time of elderly drivers and young drivers, for running scenarios that controlled the time periods (i.e., daytime or nighttime). However, we set the condition that no overtaking vehicle occurs in any case. We show reaction time (Z1) for "cognitions" in Figure 8, reaction time (Z2) for "judgments" in Figure 9, and reaction time (Z3) for "operations" in Figure 10 separately. In these figures, (a) shows the result of daytime, and (b) shows the result of nighttime. Focusing on the reaction time (Z1) pertaining to "recognition" in Figure 8, when the time period is daytime, the reaction time of the elderly drivers is significantly longer than that of the young drivers. On the other hand, when the time period is nighttime, the difference between the elderly drivers and the young drivers is hardly noticed. Generally, the visibility at nighttime is considered to be deteriorated. However, according to the questionnaire survey after the running test, the majority accounted for the elderly drivers as "the nighttime course was easier to recognize the falling objects". The reason for this trend may be attributed to the factor that eye-gaze targets during nighttime driving were limited. In the reaction time (Z2) related to "judgment" in Figure. 9, we can confirm that the reaction time of elderly drivers is significantly shorter than that of young drivers in both time periods (i.e., daytime and nighttime). This is because the elderly drivers are rich in driving experience compared to the young drivers, so we can infer that this trend is due to differences in their experiences. In the reaction time (Z3) related to "operation" in Figure 10, almost no difference is observed when the time period is daytime. On the other hand, it can be confirmed that there is a large variation in reaction time of the elderly drivers in nighttime. Comprehensively, analyzing the above results,



(a) Daytime / no overtaking vehicles

(b) Nighttime / overtaking vehicles available

Figure 10. Reaction time for operations (Z3)

we confirmed that the reaction time (Z1) related to "cognition" of the elderly drivers is longer and the reaction time (Z2) related to "judgment" is shorter than those of the young drivers.

Finally, the distance from the reference point to the falling object was calculated based on the reaction time (Z1) related to "cognition" and the speed at the time of passing through the reference point. The results are shown in Figure 11. In this section, we calculated the moving distance for all

subjects assuming constant linear motion. The clear difference can be confirmed as compared with the reaction time in Figure 8. As we can understand from Figure 11, among a series of driving behaviors, the driving characteristics related to "cognition" are greatly different between the young drivers and the elderly drivers. Consequently, we infer that the difference will greatly affect "judgment" and "operation" thereafter. Particularly, the elderly drivers have a high ability to judge based on their

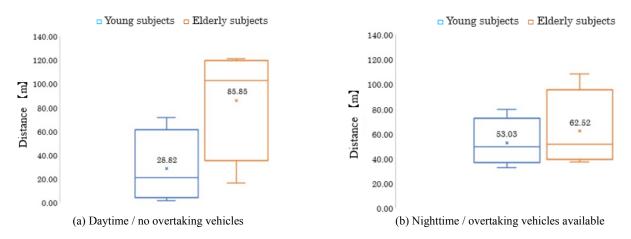


Figure 11. Distance from the reference point to the falling object.

experiences. The following framework is effective for elderly drivers, i.e., which will refine the gazing target necessary for safety driving appropriately, and direct caution resources to the gazing target at the same time.

V. CONCLUSION AND FUTURE WORK

For this study, we defined four situations that are likely to trigger accidents on highways as driving topics, i.e., breakdown vehicles, sudden appearance of small animals, falling objects, and lane decrease. Carrying out running experiments controlling the time period (daytime or nighttime) and the traffic flow (overtaking vehicle: Yes or No), we studied the comparative analysis for each driving topic, focusing on driving characteristics of young drivers and elderly drivers concerning to "cognitions", "judgments", and "operations". Results clarified the following points.

- The elderly drivers had longer reaction times than young drivers. Particularly, the reaction time regarding to "cognitions" was remarkable.
- Based on the average of whole subjects, the existence of categories characterizing elderly drivers was confirmed from the relationship between the reaction time of each driver and driving style.
- For elderly drivers, it is effective to narrow down gazing targets necessary for safety driving and allocate appropriate attention resources.

In future work, we aim to evaluate the statistical superiority of the findings obtained in this research and aim to improve the accuracy of a dangerous driving prediction model.

ACKNOWLEDGMENT

This work was supported by technical research grant of East Nippon Expressway Company Limited.

REFERENCES

- "Characteristics of Traffic Death Accidents and Criminal Investigation on Road Traffic Law Violation," Japan National Police Agency, Available from URL: https://www.npa.go.jp/publications/statistics/koutsuu/index.html [retrieved: April, 2019]
- [2] K. Kamiji and N. Takahashi, "Achievement and Effort of Safety Traffic Society for Elderly Drivers by JAMA," International Association of Traffic and Safety Sciences, vol. 35, no. 3, pp. 221-227, 2011.
- [3] H. Suzuki, "Motivating Senior Drivers toward Traffic Safety: the Traffic Sociology Viewpoint," International Association of Traffic and Safety Sciences, vol. 35, no. 3, pp. 194-202, 2011.
- [4] Y. Nakano, T. Kojima, H. Kawanaka, and K. Oguri, "Improvement of Elderly Driver's Ability by Pointing and Calling Method," The Institute of Electronics, Information and Communication Engineers D, vol. J97-D, no.1, pp. 135-144, 2014.
- [5] M. Takahara, M. Kokubun, T. Wada, and S. Doi, "Effects of a Stopping-situation Assistance System for Elderly Drivers," International Association of Traffic and Safety Sciences, vol. 36, no. 1, pp. 6-13, 2011.
- [6] K. Iida, "Hypothesis Construction on Reverse Running Occurrence Process on Highway," Grant Research by Takata Foundation ISSN2185-8950, Available from URL: http://www.takatafound.or.jp/support/articles/pdf/150626_10.pdf [retrieved: April, 2019]
- [7] M. Ishibashi, M. Okuwa, S. Doi, and M. Akamatsu, "HQL Driving Style Questionnaire: DSQ," Research Institute of Human Engineering for Quality Life, 2003.
- [8] M. Ishibashi, M. Okuwa, S. Doi, and M. Akamatsu, "HQL Workload Sensitivity Questionnaire: WSQ," Research Institute of Human Engineering for Quality Life, 2003.
- [9] G. Abe, K. Kikuchi, R. Iwaki, and T. Fujii, "Effects of Cognitive Distraction on Driver's Visual Attention," The Society of Mechanical Engineers (JSME), Transactions of the JSME (c), Vol. 76, No. 767, pp. 1662-1668, 2010.
- [10] R. Honma, G. Abe, and K. Kikuchi, "Characteristics of Visual Attention while Driving under the State of Drowsiness," Society of Automotive Engineers of Japan, Inc. (JSAE), Vol. 42, No. 5, pp. 1217-1222, 2011.

Survey on Smart Web Crawler Algorithm Based on Semantic Search Engine

Muneer Bani Yassein, Ismail Hmeidi, Yaser Khamayseh, Wail Mardini, Heba Al-Jarrah Jordan University of Science and Technology, Irbid, Jordan {Masadeh, hmeidi, yaser, mardini}@just.edu.jo, hebaatta96@gmail.com

Abstract—With the increasing amount of information and inflation of pages in the World Wide Web, it becomes difficult for the user to retrieve relevant information it needs. In other words, it leads to a big challenge for a search engine, on how to extract the exact results that match with the user query. Hence, we need smart programs such as Web Crawlers to overcome this problem. Web crawlers use different kinds of algorithms, some of these algorithms using the heuristic function to achieve the goal in an efficient manner such as A*, Adaptive A*, Bestsearch algorithm, and some other algorithms are not using a heuristic function such as Depth-first search and Breadth-first search. This survey covers these two types of algorithms and focuses on the mechanisms and the difference between them.

Keywords-Web Crawling Algorithms; Uniform Resource Locator (URL); Best-first search; Heuristic function; Web Crawler.

I. INTRODUCTION

Web crawlers, also known as web spiders and web robots, are programs that visit websites and read and download their pages and other information to create entries for a search engine index. The crawler crawls with a list of Uniform Resource Locators (URLs) to visit, called the seeds. As the crawler visits these URLs, it identifies all the URL hyperlinks in the page and adds them to the list of the visited URLs (queue), called the crawl frontier. In its simplest form, the crawler starts from the seed page and then uses external links inside it to access other pages. The process is repeated with new pages that offer more external links to follow until a sufficient number of pages is identified, or a higher-level goal is reached Web crawlers are an essential component of web search engines, where they are used to gather the corpus of web pages indexed by the search engine. Moreover, they are used in many other applications that process large numbers of web pages, such as web data mining, and comparison-shopping engines. [1]. The crawler can start from any seed URL, but this is not enough to reach all the web pages. The pages referenced by the seed URLs should not reference it back to them, otherwise it will end up in an infinite loop, and this is a crucial factor to be considered. Hence, it is essential to start from a proper seed URL. For example, Bing or Google

Dragana Krstic Faculty of Electronic Engineering, University of Niš, Niš, Serbia email:dragana.krstic@elfak.ni.ac.rs

search engines can be used to get seed URL by merely entering the keywords into them and consider their resulting links as seed URLs. It is not possible to cover and index all the websites in the World Wide Web for a particular search entry. A web crawler always downloads a fraction of the web pages that contain the most relevant pages and not just a random sample of the Web [2] [3]. The relevance of a page depends on some factors such as the number of visits on the page. Web crawlers employ different strategies for selecting the websites to be downloaded, such as depth-first, breadth-first, best first, A* and other algorithms that will be covered in this survey.

The crawler aims to crawl the World Wide Web and report back some essential data and then perform an initial data analysis using some additional data before it stores the collected data. To achieve this purpose, the crawler that crawls the Internet must have the following basic features to serve their purpose [5]:

a) Robustness: The Web contains loops which are aims to mislead the crawler to recursively crawl a particular domain and get stuck in one single domain.

b) Distributed: The crawler should be able to work in a distributed manner to crawl the Internet as quickly as possible.

c) Scalable: The crawler should have the flexibility to add new machines whenever necessary.

d) Extensible: The crawlers should be able to adapt to the increasing amount of data formats on the Web.

e) Quality: The crawler should be able to distinguish between useful and useless information. In other words, they should be able to filter out the relevant content.

f) Freshness: The crawler should make sure that the concepts on the search engine are the latest and relevant to the present context.

This paper is organized like this: in Section 2, the architecture of web crawler is presented, and different types of web crawlers are listed. The web crawler strategies are specified in Section 3, and conclusion and future work are shown in the last, fourth section.

II. ARCHITECTURE AND TYPES OF WEB CRAWLERS

The architecture in Fig. 1 [6] provides an outline of the main modules of a web crawler. URL frontier contains a list of visited URLs (queue) to be fetched. A DNS resolution module converts the URLs to its equivalent IP addresses. Fetch module that uses the HTTP request to retrieve web pages associated with its URL. A parsing module extracts the query text and set of links from a fetched web page. Finally, a duplicate elimination module checks to determine whether an extracted link is already in the URL frontier or has recently been fetched.

There are different types of crawler; in this survey, some of them will be covered.

A. Focused Crawler

This kind of crawlers works by gathering documents that have specific areas of interest. In other words, it collects the documents similar to each other. The main thing to be kept in mind is that the page is downloaded after it is estimated that the page is similar and relevant to the given topic. The main advantage of this type is the availability of the URLs without downloading the page to predicate the similarity of not traversal page, meaning fewer costs.

B. Distributed Crawler

This type of crawlers distributes the work to several other crawlers to perform the task at the same time. Then, it collects all downloaded pages from all crawlers and sends them to a central indexer where links are extracted and sent via the URL server to the crawlers. The main advantage of distributed crawler is that it is strong. It can withstand system crashes and other types of problems and can adapt to different crawling requirements.

C. Parallel Crawler Algorithm

Parallel Crawler (PC), which means more than one crawler, runs multiple processes in parallel, where PC depends on freshness and Selection Page [7]. The main idea for PC is to maximize the download rate and achieve minimum overhead by using parallelization. We can find more than one crawling processes for the same URL. PC can be distributed in different places or at the same local network [8]. In [7] and [8], the authors presented different algorithms that will help to achieve high performance.

III. WEB CRAWLER STRATEGIES

In this section, several search algorithms used by Web crawlers will be enumerated and described.

A. Breadth First Search

It is the simplest form of crawling algorithms, which starts with the root URL links and searches on the connected links URL (in the same level). This algorithm is also known as Blind search algorithm. That is because it is not considering any information about the topic and means the breadth-first search requires the maximum amount of searching time.

On each page the crawler visits, all external links (hyperlinks) will be extracted from the web pages, and all these links will be distributed in the searching tree on the leaves of the current node [9].

In Fig. 2, we show an example where the search starts from the root URL (i.e., Seed page 1) and then collects URL page 1.1 and page 1.2 which get searched in sequence at the same level, then pages 1.1.1, 1.1.2, 1.2.1 and 1.2.2 are searched in order [9].

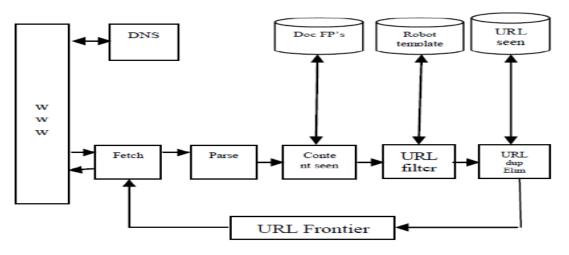


Figure 1. The architecture of crawler.

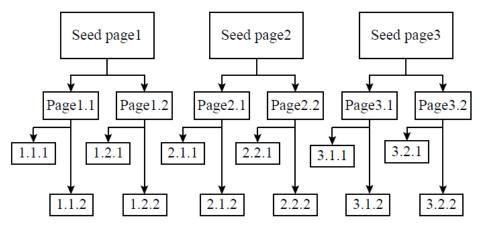


Figure 2. Web crawler forest.

B. Depth First Search

In this search algorithm, the crawler starts from the root URL and traverses to child URLs in the deep form. If there is more than one child, the priority will be going to the most left child and then keep performing deep traversal until it receives no more nodes for searching steps. After that, backtrack to the previous node to check whether there exists another unvisited node, and then continue similarly. This algorithm might end up in an infinite loop in case of large number of branches.

In the example in Fig. 2, the search starts with root URL (i.e., Seed page 1), and then the search proceeds by downloading pages 1.1, and 1.1.1, then page 1.2 and its own child's are searched in order.

C. Best First Search

Best First Search is an informed search algorithm based on estimate function which predicts the similarity between the candidate URLs to pick the best one to fetch. This estimated value could be the score or rank of relevancy. There are many algorithms to apply the classifier for these heuristic values, such as the cosine similarity, Support Vector Machine (SVM) and string matching used for scoring [10].

Unlike Uninformed search, the best first search algorithm uses heuristic function to improve its result as much as possible by exploring the node with the best score first [5].

D. A*Search

A* algorithm uses the best first search algorithm to estimates the lowest total cost of any solution path. It combines the features of uniform-cost search and greedy (heuristic) search to compute the optimal path solution. The cost associated with a node is f(n) = g(n) + h(n), where g(n) is the cost of the path from the initial state to the goal node, and h(n) is the heuristic estimate for the cost of the path from node n to the goal node. f(n) can be referred to as the relevancy cost of each link.

The A* algorithm gives the best result of relevant retrieve information with the lowest time compared with other algorithms that don't use a heuristic function.

E. Adaptive A* Search

Adaptive A* Search is an informed heuristic search. In each iteration, it improves the relevancy value of the Web page and uses this new update in the next iteration. At the beginning of the process, it takes more time because it stores the history of previous iterations with every search. Once it gathers sufficient knowledge about the relevant pages, it needs the same time as A* search, Best-first search and any other heuristic-based search algorithms.

F. Page Rank Search

Page Rank algorithm works by counting the number and quality of links to a page in order to determine how important the website is. So, as it is shown by the equation below, the Page Rank does not rank websites but is determined for each page individually

$$PR(P_1) = PR(T_1)/L(T_1) + \dots + PR(T_i)/L(T_i).$$

To find the Page Rank for a page, called $PR(P_1)$, we need to find Page Rank of all pages T_i with inbound links and outbound links to page P_1 . For example, page T_1 , which has a link from P_1 , is calculated. Then page $L(T_1)$ will give the number of outbound links to the page P_1 . We do the same for T_2 , T_3 , up to page T_i , then the Sum of all the weighted Page Ranks of all pages is calculated.

G. Path-Ascending Crawling Algorithm (PAC)

The crawler tries to obtain more information for a specific website by crawling every path for that URL [11]. For example, in Fig. 3, we have a URL: http://llama.org/hamster/monkey/page.html, PAC will try to crawl /hamster/monkey/, /hamster/, and /. The added values for PCA algorithm become very efficient while finding the resource.

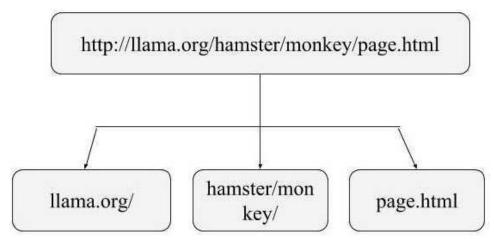


Figure 3. Path-Ascending Crawling Algorithm (PAC)

TABLE I. COMPARISON BETWEEN ALGORITHM

| Algorithm indicator | Number of the relevant pages | Time to retrieve |
|----------------------|------------------------------|------------------|
| Breadth-first search | Extremely small | Extremely large |
| Depth-first search | Quite small | Large |
| Best-first search | Large | Low |
| A* algorithm | Extremely large | Extremely Low |
| Adaptive A* search | Large | Low |
| Page Rank search | Large | Large |

IV. CONCLUSION AND FUTURE WORK

Reducing the search time with the best relevant result is one of the main issues being targeted by search engines these days. This survey focuses on the Web crawling algorithms that use heuristic function and do not to retrieve relevant data. Breadth-first search is a blind algorithm and is not an efficient method. The depth-first search may end up in an infinite loop when the branches are too large. The Best first search and A* algorithm show nearly the same result with respect to search time. They can be improved by enhancing the heuristic function. Adaptive A* performs more efficiently when the users search at the same type of content because it depends on the history of searches, and with each search, the efficiency of the search will increase. This may consume some amount of time compared with the A* algorithm. All algorithms discussed in this paper are effective for search engine, however, the advantages favor algorithms that use the heuristic functions to retrieve the best relevant URLs and least response time.

REFERENCES

- [1] M. Najork, Web Crawler Architecture, Microsoft Research, Mountain View, CA, USA
- [2] V. Kancherla, A Smart Web Crawler for a Concept Based Semantic Search Engine, San Jose State University, 2014.

- [3] D. Yadav, A. Sharma, and J. P. Gupta, "Parallel crawler architecture and Web page change detection", WSEAS Transactions on Computers, issue 7, volume 7, July 2008, pp. 929-940.
- [4] S. M. Pavalam, S. V. Kasmir Raja, M. Jawahar, and F. K. Akorli, "Web crawler in mobile systems", International Journal of Machine Learning and Computing, International Journal of Machine Learning and Computing, vol. 2, no. 4, August 2012. pp. 531-534.
- [5] C. D. Manning, P. Raghavan, and H. Schütze, Introduction to Information Retrieval, Cambridge University Press. 2008.
- [6] R. Thite, B. Pawar, T. Mode, and M. Mete "Survey paper on smart Web crawler", International Journal of Novel Research in Computer Science and Software Engineering, vol. 3, issue 1, pp. 274-277.
- [7] S. Chakrabarti, M. Van den Berg, and B. Dom, "Focused crawling: a new approach to topic-specific Web resource discovery", Computer Networks, 31.11-16, 1999, pp. 1623-1640.
- [8] N. Tyagi and D. Gupta. "A novel architecture for domain specific parallel crawler", Department of Computer Engineering, Shobhit University, Meerut, India, 2010.
- [9] Y. Yu, S. Huang, N. Tashi, H. Zhang, F. Lei, and L.Wu, "A survey about algorithms utilized by focused Web crawler", Journal of Electronic Science and Technology, vol. 16, no. 2, June 2018.
- [10] R. Dechter and Judea Pearl. "Generalized best-first search strategies and the optimality of A*", Journal of the ACM, 32(3), pp. 505–536, 1985, doi:10.1145/3828.3830
- [11] V. Cothey, "Web-crawling reliability", Journal of the American Society for Information Science and Technology 55.14, 2004, pp. 1228-1238.
- [12] R. Janbandhu, P. Prashant, and M. M. Raghuwanshi, "Analysis of web crawling algorithms", International Journal on Recent and Innovation Trends in Computing and Communication, 2.3, pp. 488-492, 2014.

Innovation as a Structural Attribute of Company Competitiveness

Oswaldo Luiz Agostinho School of Mechanical Engineering – FEM, University of Campinas , UNICAMP Campinas, SP, Brazil e-mail : agostinh@fem.unicamp.br

Abstract-Innovation does not occur spontaneously in an organization. Instead, structural conditions providing methodological support to the innovation process must be established, thereby ensuring the competitive edge required to respond to market demands. Organizations should therefore be designed with attributes, i.e., internal characteristics that provide structural conditions aimed at meeting these external demands. This paper aims to characterize the innovation process as one of the structural attributes of company competitiveness. Innovation, which is considered an attribute aimed at meeting customer needs, is shown to be a prerequisite for organizations to gain a competitive edge in the market. The methodology adopted here consists of identifying the structural attributes of the organization, being innovation one of them, establishing the conditions required for the innovation process to prosper. Sub attributes that support the innovation attributes are presented and linked to the innovation process. The main result is a model that considers the cause-and-effect relationship between external demands and the structural attributes of innovation, providing methodological support for innovation processes, enabling companies to meet the needs of competitiveness in their market segment. This model facilitates the design of innovation characteristics, creating guidelines for the company's strategic positioning in its market segment. This approach, linking concepts of competiveness and innovation, is not usual in the existing literature and business practices . The limitation of this research is mainly the design of this generic model for the specific organization application.

Keywords-Innovation; organization attributes; competitiveness; business models.

I. INTRODUCTION

Competitiveness is a critical success factor for organizational survival in the turbulent 21st century. Competitiveness can be understood as a company's ability to continuously review its competition strategies in order to gain a favorable position in its market segment. It is triggered in response to external pressures resulting from the demands of market, science and technology and society, which are the main actors of the external environment. These demands are expressed through continuous and intermittent stimuli produced by these actors, e.g., shorter product life cycles, increasing product diversification, reduction of time to market, rapid obsolescence of existing products and organizations, swift emergence of new products, and the need to update the knowledge base. These factors exert different levels of influence, according to their geographic location (county or municipality, state, region, country, or continent) and economic activity (health, education, durable or non-durable consumer goods, capital goods, etc.) [3][10][12]

Company competitiveness means that the organization's working conditions enable it to generate higher than average profits in its market segment, operating sustainably, with quality, speed and flexibility. Moreover, the organization must satisfy stakeholders and comply with environmental requirements. Competiveness can be divided into External competiveness, which means the company has the ability to persuade consumers to buy its products rather than those of its competitors - in other words, the ability to shift the consumer's brand loyalty. External competitiveness is calibrated and controlled by qualitative standards. This ability, in turn, drives so-called Structural competitiveness, which is seen as a set of harmonious and synergic methodological and technological factors that induce the external market to shift its product loyalty from the brand it previously purchased to the new one. Structural competitiveness, which is calibrated and controlled by quantitative standards, is achieved through the use of structural attributes. Figure 1 shows a schematic diagram of the competitiveness model, identifying the actions of each of its components and their respective relationships [1].

This paper is structured with the following sequence :I-Introduction: This section presents the more wide concepts like Competitiveness and its influence in the organization structure ;II - Structural attributes presents the concept of Structural competitiveness, its quantitative standards and relationship with qualitative standards, the and characteristic of the external competitiveness main attributes; III - Innovation : this section provides the Innovation definitions , its classification and the its respective sub attributes, information flow needed for innovation; IV-Relationship between innovation and other structural attributes: this section establish the relationship between innovation and the other structural attributes as Organization driven attributes and Human capital driven attributes; Conclusion and future work: this section presents the main conclusions and the proposed future works.

II. STRUCTURAL ATTRIBUTES

This section presents the competitiveness model, showing the concepts in the external environment and the organization environment; it also presents the relationship between the external and organization environment.

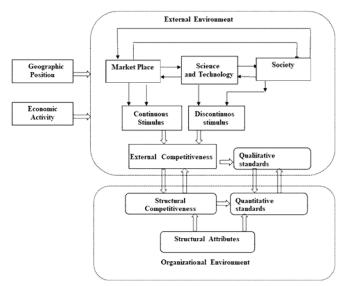


Figure 1.Competitiveness model

Companies achieve structural competiveness as a result of their market, organizational and human capital management structures. These are defined as *structural attributes*, which are characteristics of the management system, manifested by ongoing practices composed of methods that make up the organizational and technological management. The attributes and subattributes of competitiveness are classified according to the range of applications described below.

A. Market driven attributes:

Characteristics inherent to the organization's management system, expressed by an ongoing, comprehensive and integrated set of methods aimed at providing the conditions required to meet consumer market demands. Market driven attributes can be divided into the following subattributes:

1) Innovation - The process of technological innovation consists of a complex set of activities that transforms ideas and scientific knowledge into physical products employed in the real world [5][12]. Innovation also comprises existing technology and inventions to create a new product, service or process. Innovative companies search for all kinds of ideas, and their organizational culture encourages the development of these ideas into feasible business programs. The subattributes of innovation can be divided as follows: a) focus on consumer needs; b) enterprise-wide quality; c) enterprise-wide planning; d) use of core knowledge; and e) continuous upgrading of products and processes.

2) Agility - An agile company is one whose technological and administrative infrastructures are flexible and can be rapidly created, set up, and rearranged to meet external business needs. Among other things, this attribute favors faster time to market of new business initiatives. The subattribute of agility can be divided into: a) technological independence; b) reuse of solutions; c) suitable business infrastructure; d) system-wide response capacity, not just of individual processes; e) existence and implementation of strategies; and f) enterprise-wide business architecture, including supply chain and marketing processes.

3) Responsiveness - An organization's ability to respond to market demands in a timely manner, also called "time to market." This ability includes proactive search for customer and supplier feedback, fast cycle organization, flexible thinking, and doing things in short periods of time. The subattributes of responsiveness can be divided into: a) business process optimization, with the elimination of activities that do not add value to the product or business; b) adoption of proven technology; c) integration of business strategies; e) fast response time from Engineering and Marketing processes; and f) supply chain network optimization.

B. Organization driven attributes

Characteristics inherent to the organization's management system, which are indicated by ongoing, comprehensive and integrated methods aimed at providing favorable conditions for its organizational structures to meet external needs. Organization driven attributes determine how the business administration and management process is applied throughout the organization. These attributes can be divided into: a) information flow synergy; b) structural integration; c) information infrastructure; and d) knowledge management.

C. Human capital driven attributes

Characteristics inherent to the organization management system expressed by a continuous, comprehensive and integrated set of methods aimed at providing favorable conditions for the development of selection, education, and training programs that enable the organization to reach the level of human capital needed to support its competitiveness. Human capital driven attributes can be divided into: a) Teamwork design; b) project clusters; c) virtual teams; d) human networking; and e) assumption-toknowledge ratio.

It can be stated that a company's structural attributes enable it to be organized, resulting in a state of structural competitiveness that should be unambiguously aligned with external competitiveness. To reinforce this mindset, the organization must have internally structured attributes whose implementation enables the company to meet the demands of external competitiveness in response to continuous and intermittent stimuli. Given the limited scope of this article, however, our discussion will focus solely on the innovation-related portion of the complete competitiveness model.

Figure 2 schematizes the attributes and subattributes of competitiveness, split in market driven, organization driven and human capital driven.

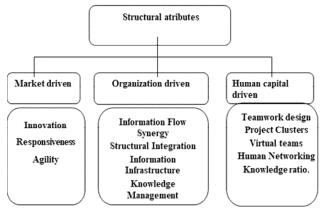


Figure 2.Structural attributes

III. INNOVATION

Any discussion about innovation must include the subject of technology and their interrelatedness [5]. Technology involves the design, production and distribution of goods and services in response to market needs. Invention, on the other hand, is limited to expressing a new idea, a new concept, or sketching a new path without considering real technical or market feasibility. Hence, innovation bridges the gap between ideas and their actual implementation to meet human needs. There is a natural link between technology and innovation. Technology is improved continuously through a flow of incremental innovations that build and shape its course.

Before discussing structural aspects, some concepts of innovation will briefly be reviewed [14]. Innovation involves the creation of a product, service or process that is new to an organization. It is introduced on the market through the use or sale of a new product, service or process. Innovation can also be defined as the creation of something new. It does not have to be new to the world; rather, it is seen as the first use of an Idea within an organization, whether or not this Idea has been adopted by other organizations. The technology (or product) must be novel or ground-breaking. An innovation may be a change in an industrial practice, which improves productivity. The innovation process involves integrating existing technology and inventions to create a new or improved product, process, or system. Although invention and innovation are closely related, they are not the same thing. An invention can be considered an event, while an innovation can be thought of as a process. Innovation is less common, while an invention usually precedes several innovations. Innovation represents the important link between an Idea and its exploitation or commercialization. The bottom line of innovation is the market, which will buy it or ignore it, thereby dictating its success or failure. The relationship between innovation and the need of competitiveness is a partially solved task. This paper intends to identify a significant, challenging, not yet solved, relationship , and propose the way to reach it

A. Classification of Innovation

Innovation can be classified as radical or revolutionary, or as incremental or evolutionary. Radical breakthrough innovations are usually based on inventions. They are relatively rare and typically start outside the boundaries of a firm. When such innovations are developed within the boundaries of a firm, they signify the introduction of something that is not only new to the organization but also drastically different from existing practices. An example of radical innovation is the substitution of chemical photographic processing for computational photography.

The second category, which comprises incremental or evolutionary innovations, involves small but important improvements in a product, process, or service. Incremental innovations are relatively common and are created in-house, helping companies maintain their competitive edge in the market. An example is the Kaizen method, a process of continuous improvement which is typical of this type of incremental innovation.

Routine innovation is another category occasionally used to refer to the introduction of something new, albeit very similar to what an organization had in the past. An example of this is the redesign of an item for a new one in the manufacturing operation of a product component.

B. Innovation as a structural market driven attribute

As Figure 2 indicates, innovation is a market driven structural attribute. This means that, for an innovation to be of interest to the business, it must be part of the concepts that underpin the company's management model and be included in its strategic issues and plans. Hence, it must be linked, via cause-effect relationships, to the various components of the business model. Considering the above described definitions and implementation of the concepts of innovation, they can be contextualized as a market attribute that is an integral part of the competitiveness model. Moreover, innovation can be classified as a structural attribute of the organization, whose aim is to satisfy customer needs.

C. Innovation subattributes

The innovation subattributes are characteristics of an organization that creates favorable structural conditions for the development of innovation processes [12][14].

1) Focus on consumer needs

Meeting the requirements of this subattribute involves changing an organizational culture that is often deeply embedded. Such changes consist of: i) shifting from an internal and often functional approach to a customeroriented matrix management system; and ii) searching for customer input about products and services and remaining constantly attentive to ensure that this input is communicated throughout the organization instead of only to marketing and Research and Development staff. Several surveys have revealed that most of the development of new ideas and products comes from listening carefully and working with clients, rather than from new ideas for developing new products created separately in Research and Development laboratories. Hence, innovation efforts should preferably be aimed at meeting customer needs.

2) Enterprise-wide quality

The cost of entry for companies in the 21st century is the highest possible quality. To this, they end, will have to adopt the broadest possible vision of quality. The concept of total customer satisfaction must be a goal. The company's entire management team must be fully committed to the zealous pursuit of this goal. This quality mindset must be applied to both the tangible and intangible services pertaining to the company's physical products. Quality has to do with how people focus on the customer, i.e., the next person in line responsible for the subsequent activity of the company's business processes (the internal customer) or the all-important end user. It is essential for customer supplier relationships to be established within each company's business process. The innovation process must be imbued with the company's quality concepts to ensure that its products and services meet customer needs and expectations with the highest possible level of quality.

3) Enterprise-wide planning

To ensure its success, the innovation process must be included in the full enterprise-wide planning at the various strategic, tactical and operational levels. In other words, the company must formulate innovation strategies, i.e., tactical plans for innovation originating from its innovation strategies, operational plans for innovation originating from its tactical plans, and specific projects originating from its operational plans. Large organizations have a natural tendency to create internal mechanisms that hinder innovation. It is a major strategic challenge to reduce as much as possible these mechanisms that favor operations, i.e., efficiency of current activities in detriment to creativity, idea generation and the launch of innovative projects. Innovation is initially seen as a disruption that interferes in the operations of existing activities. In traditional and conservative organizations, the difficulties arising from the introduction of a new software program (e.g., an ERP system), a new method (Group technology, Simultaneous Engineering), or a new organization model (Business Process Management), etc., are well known. Such

organizations were designed to last. Rational decisionmaking and operational pressures depend on many layers of reporting and control. This obviously tends to favor the amortization of heavy investments in existing activities. In this regard, organization and innovation are on opposite sides of managerial concerns. Companies need not only to operate but also to evolve. Hence, the inclusion of innovation processes among the company's issues, and thus, of innovation strategies in its long range strategic plan, is a high level decision that is usually made by the Board of Directors and conveyed as a top-down decision. Therefore, the company must be convinced, at the political and strategic level, that investing in innovation is a business decision resulting from the realization that it has to maintain and exceed its current competitive edge in order to satisfy customer needs and expectations.

4) Use of core knowledge

Regardless of their organizational structure, considerable duplication of efforts and management fragmentation are common issues in companies. In large companies, in particular, opportunities to achieve operational synergy are rare, since these companies tend to focus on purely financial aspects. Thus, it is essential to define core competencies, representing unique corporate capabilities, i.e., the company's "collective knowledge." Companies normally have only two or three core competencies. These core competencies and expertise should be defined by the strategic positioning of the company, which should allocate the corresponding funds. Moreover, top management should use the company's core knowledge as basis not only to build strategic alliances, but also to create sourcing and vertical integration strategies. Clearly, organizations must offer innovative products and services if they operate in external environments where product lifecycles decrease continuously and which are characterized by ever increasing product diversification and by the demand for products with highly differentiated functionalities and performance. In such organizations, innovation is essential for them to maintain their competitive advantage. Therefore, innovation must be part of the "collective learning" process, which is considered a core competency.

5) Continuous upgrading of products and processes

The process of ongoing improvement and of adapting to the environment of continuous change, ideas and plans that make up for continuous upgrading of processes and products is a constant demand. In this context, incremental and routine innovations in products and manufacturing processes, and the elimination of activities that add no value to the company's business processes are some examples of these innovative practices.

D. Innovation information Flow

The diagram in Figure 3 illustrates the company's response to external demands for new products and services, and the reaction of its innovation process with new products and services. The external flow illustrates how the external environment activates the organization, calling for new products and/or services. Enterprise flow takes place in the opposite direction, reacting through the innovation flow or process by supplying the products or services required by society, science and technology or the market.

The attribute and subattributes of innovation, considered a market driven structural attribute, uses the external needs as input and develops new products and services, following the connections between its subattributes and other structural attributes, such as organization and human capital. These relationships between innovation and other structural attributes are discussed below. Figure 3 shows the information flow between the external environment and structural attributes, detailing this flow in the innovation process.

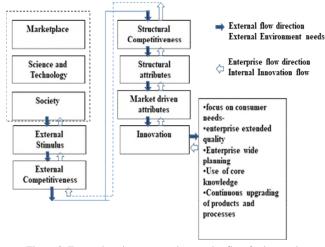


Figure 3. External environment and enterprise flow for innovation processes

IV. RELATIONSHIP BETWEEN INNOVATION AND OTHER STRUCTURAL ATTRIBUTES

The attribute of innovation depicted in Figure 2 is correlated with other structural attributes, as follows:

A. Organization driven attributes

Organization driven attributes determine how the business administration and management process is applied throughout the organization. The innovation process is supported by these attributes, so that its concepts and process activities can be applied continuously. Organization driven attributes are divided into the following subattributes. [1][12]

1) Information flow synergy

The company's business processes, including its innovation process, should involve activities that add as much value as possible to the business.

2) Information infrastructure

The existence of a unified information system that includes most of the company's business processes facilitates interactions between its innovation process and other business processes, such as product design, product manufacturing, marketing and sales, manufacturing planning and control, etc.

3) Knowledge management

The innovation process will be facilitated if is linked to the knowledge base required for selected technologies, internal developments, R&D partnerships or acquisitions.

4) Structural integration

The structural integration of the organization's business processes enables the innovation process to be involved in the management of research activities, development feasibility studies, and more generally, in project management. Therefore, structural integration becomes part of the implementation and improvement of subsequent product and process technologies integrated in the company's portfolio.

B. Human capital driven attributes

Like the company's other processes, the innovation process is supported by the skills of human capital[17]. Ultimately, it is people that distinguish successful companies from those that fail to provide the means for their human resources to develop. People create and implement strategies, design information systems and make them work. Hence, it is obviously people that create inventions and develop innovations. However, for people to be differentiators, they must be educated and trained in a continuous learning environment. Education is only part of the solution. People can also be encouraged by performance related ideas and metrics, and empowered to get the job done. The creation of an innovative environment must follow the abovementioned parameters, with the added premise that an innovative product or service that fails when initially introduced on the market should not be considered a failure, but rather, an opportunity to review ideas and designs and use these reviews in the next innovative project. The structural sub attributes driven by human capital are as follows.

1) Teamwork design

The innovation process is facilitated and accelerated when people involved in innovation engage in teamwork, exchanging ideas and concepts without putting up selfprotective barriers. Teamwork design is the opposite of peer-to-peer organization, which is a frequent feature in traditional organizations. The innovation process may be slower in peer-to-peer organizations because the top down management style often hinders the free expression of creative minds.

2) Virtual organization and virtual teams

These subattributes are feasible when the information technology infrastructure enables knowledge to be exchanged through available networks, and innovation groups are formed in organizations located in geographically different regions or countries within the same extended enterprise [5][17].

Figure 4 illustrates the main relationships between innovation and the organization driven and human capital driven attributes. It shows the relationship between external environment, structural attributes, innovation and the organization driven and human driven attributes components.

V. CONCLUSION AND FUTURE WORK

Organizational competitiveness is essential for companies operating in their different fields of activity. In the turbulent environment of the 21st century, with product life cycles decreasing and product diversification increasing continuously, the demand for innovative products and services is growing apace.

Innovation does not come about spontaneously in organizations. In fact, the innovation process requires methodological support provided by structural conditions, which in turn must be tied to the competitive requirements manifested by external markets. This article discusses innovation as a process rather than as an event. Hence, it should be seen as an organizational process and included as a structural market driven attribute in the competitiveness model. In this proposal, the sub attributes of the innovation process are presented as the basis for its development and ongoing feasibility. The logical connection of innovation sub attributes to the organizational and human capital driven, internal attributes of competitiveness (Figure 4), and actors of the external environment (Figure 3) is considered understandable and feasible solution to connect Innovation with external Competitiveness, as proposed in Figure 3. Nevertheless, this process should be supported by other organization driven and capital driven structural attributes as well. The various workflows that comprise the innovation process as part of the competitiveness model are presented in this paper by means of cause-effect relationships between external requirements and the structural attributes of innovation. As proposed, a logical relationship was developed between the innovation process as market driven attribute and its competitiveness, necessary to the company survival in the third millennium. This approach is presented through a model that considers the cause-and-effect relationship between external demands and the structural attributes of innovation.

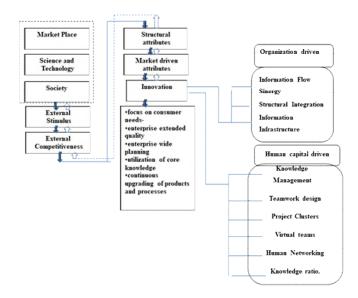


Figure 4. Main relationships between innovation and organization driven and human capital driven attributes.

Lastly, in view of the above described concepts, it can be stated that it is advantageous for an organization to include the innovation process in the structural characteristics of its management system, provided this process meets external needs, thereby boosting the company's competitive edge.

As future work, we are planning to address the determination relationships following: a) between innovation and organization driven and human capital driven attributes that are more relevant to competitiveness that are preponderant to reach the necessary competitiveness level; b) use focus methodologies to establish these prevailing attributes. The lessons learned that in this article are: innovation cannot be seen as an independent characteristic of an organization or company, but as a market driven attribute that contribute to the company to reach the competiveness level; it is important to carefully establish its relationships with other characteristics and attributes of the enterprise that are not clear at first view.

ACKNOWLEDGMENT

This article was developed based on a series of surveys conducted by the author and his students at the University of Campinas, SP, Brazil. These surveys involved several Brazilian and multinational companies operating in the region of São Paulo. The application of the competitiveness model, and particularly, the innovation processes was emphasized in the students' master's and doctoral thesis. The author would like to thank all these students.

REFERENCES

- [1] O.L.Agostinho,"Sistemas de Fabricação",Editora da Universidade de São Paulo, 2018
- [2] O.L.Agostinho,"Engenharia de Fabricação Mecanica," Editora Elsevier, 2018
- [3] O.L.Agostinho, "Business Processes as basis of organization framework to support organization driven organization", ECIME, "7th European Conference on Information Management and Evaluation, 2014
- [4] F.Betz,"Managing Technological Innovation", John Wiley and Sons, Inc. New Jersey, 2009
- [5] C.M.Christensen,"The Innovator's Dilemma",Harvard Business School Press . Boston , USA,1997
- [6] M.A.Cusumano,"Thinking Beyond Lean", The Free Press; New York, 2005.
- J.J. Donovan, "The Second Industrial Revolution; Business Strategy and Internet Technology", Cambridge Technology Group, 2013
- [8] Harvard Business Review on Knowledge Management ,The Harvard Business School Press ; Boston , USA , 2004
- [9] R.McDowell and W.LSimon," In Search of Business Value: Ensuring a Return on Your Technology Investment", Select Books Inc., New York, New York, 2007
- [10] European Institute for Technology and Innovation Management ,"Bringing Technology and Innovation into the Boardroom", Palgrave Macmillan, New York, USA, 2006
- [11] R.C. Gomes, and S.V.R.Galina,"Inter Organizational Innovation Networks of Brazilian and Spanish Biotechnology Companies". International Journal of Engineering Business Management, vol. 9., 2017
- [12] T.G. Gunn ,"21st Century Manufacturing", Harper Collins . New York , USA.- 2004
- [13] J.Jeston, and J.Nelis,"Business Process Management" ,Elsevier Ltd, London, 2006.
- [14] T.Khalil," Management of Technology", Mc Graw Hill , New York, USA. (2000
- [15] M.Rivastava and A.Franklin,"Building a Sustainable Competitive Advantage,",Journal of Technology Management and Innovaton,LLC. San Francisco, CA 95050, USA.- 2013
- [16] I.S.Rampazzo and L.H.Mendonça, "Selection of Methodologies to be applied in the market driven attribute of Competitiveness" (in portughese) – University of Campinas, São Paulo, Brasil.
- [17] C.M.Savage,"5th Generation Management",Butterworth-Heinemann. Boston, USA-2009
- [18] R.J.Schonberger," World Class Manufacturing; The Next Decade", The Free Press, New York, 1999
- [19] T.Toivonen," Continuous Innovation, Combining Toyota Kata and TRIZ for sustained innovation", TRIZ FUTURE, TF 2011-2014, Procedia Engineering.