ITHEA

International Journal INFORMATION THEORIES & APPLICATIONS

2015 Volume 22 Number 1

International Journal

INFORMATION THEORIES & APPLICATIONS

Volume 22 / 2015, Number 1

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

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

Responsibility for papers published in IJ ITA belongs to authors.
General Sponsor of IJ ITA is the **Consortium FOI Bulgaria** (*www.foibg.com*).

International Journal “INFORMATION THEORIES & APPLICATIONS” Vol. 22, Number 1, 2015

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

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

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

CHARACTERIZATION OF A CUFF-BASED SHAPE MEMORY ALLOY (SMA) ACTUATOR

Gang Wang, Michael D. Poscente, Simon S. Park,
Orly Yadid-Pecht, Martin P. Mintchev

Abstract: *This article proposes a novel cuff-based microactuator for automatic blood extraction using Shape Memory Alloy (SMA) technology. It aims to provide an actuator solution for an electronic mosquito-like device, e-Mosquito, which enhances blood glucose measurement to off the potential to significantly improve both the quality of life for diabetic patients and their disease management. Fabrication and assembly methods of the microactuator prototype are discussed. The 6-mm thick microactuator prototype underwent mechanical tests in a laboratory environment to characterize its penetration force and depth, which are the major factors facilitating skin penetration and blood extraction. Testing results demonstrated that this actuator can produce a maximum penetration force of 160gf and a maximum displacement depth of 2.45mm, which both exceed the minimum force and depth requirements for blood capillary-reaching skin penetration in humans. Five actuator prototypes were assembled and integrated to a cuff wrist strap. A pilot human study was performed to test the blood extraction capabilities of the device. All five actuators successfully penetrated the skin and drew whole blood samples appearing on the skin surface of the subject.*

Keywords: *BioMEMS, Shape Memory Alloy Actuator, Diabetes, Blood Extraction.*

ACM Classification Keywords: *J.3 Life and Medical Sciences.*

1. Introduction

1.1. Diabetes Mellitus

Diabetes Mellitus is a systemic disorder that results in elevated blood glucose levels due to insulin deficiency in the body, which can lead to many secondary complications [ADA, 2012a]. The disease affects over 400 million people worldwide and the numbers are continuously climbing. Type 1 Diabetes Mellitus (T1DM) refers to an absolute insulin deficiency due to autoimmune destruction of islet cells in the pancreas [ADA, 2012b]. The discovery of insulin by Banting in 1921 became a treatment for T1DM, which at that time was considered untreatable. Presently, T1DM patients account for about 10% of the total diabetic population and are insulin-dependent throughout their lifetime. Type 2 Diabetes Mellitus (T2DM) includes individuals who have insulin resistance but incomplete insulin deficiency. T2DM accounts for about 90% of the total diabetic population [ADA, 2012a]. Of these patients, about 20% require external insulin-based maintenance similar to the one for T1DM patients. Over time, high blood glucose levels can lead to many diseases including kidney failure, nerve damage and blindness [CDACPGEC, 2013]. Both Type 1 and Type 2 Diabetes Mellitus require long-term maintenance, the goal of which is to achieve optimal glucose monitoring and control with the aim of decreasing the risk of vascular complications while minimizing daily glycemic variations [Hendriks et al., 2000].

1.2. Available FDA-Approved Glucose Monitoring Techniques

1.2.1. Fingerpricking Test

Standard blood glucose monitoring for diabetics relies on the fingerpricking test, a procedure in which a finger at its tip is pricked with a needle in order to obtain tiny amount of capillary blood for glucose testing by a sensor on a pen-like device. Although it is highly accurate for detecting blood glucose levels, fingerpricking test is painful and inconvenient. Therefore, patients, especially those who have developed visual impairment caused by diabetes (diabetic retinopathy), are often unable to adhere to the test schedule. As a result, the discontinuities in glucose monitoring limit the applicability of the fingerpricking test as a reliable tool in the intensive management of diabetes [Penfornis, 2011].

1.2.2. Continuous Glucose Monitoring

With the advances in implantable microelectronics, wearable devices featuring real-time continuous glucose monitoring (CGM) started to emerge on the market in the late 1990s [Penfornis, 2011]. CGM devices do not measure blood glucose levels. Instead, glucose levels in the interstitial fluid (ISF) are monitored by a needle-type sensor implanted subcutaneously. In the past ten years several companies released various CGM products, which include Abbott FreeStyle Navigator®, MiniMed Paradigm®, MiniMed Guardian® and DexCom SEVEN® PLUS. CGM is less invasive than the fingerpricking test. However, its accuracy is dependent on the equilibrium of glucose levels between ISF and whole blood. The balance between the two glucose levels further accounts for a time delay in the measurement and requires frequent recalibration using fingerpricking tests [Hoeks et al, 2011]. The price of achieving reduced invasiveness is the decrease in measurement accuracy. In particular, the false positive rate increases significantly due to sweating, temperature changes, electrostatic noise sources, etc., not to mention the invasiveness of the implantation of the ISF glucose sensor and the possibilities for infections [Hoeks et al., 2011]. For this reason, FDA approved the use of CGM as a supplementary technology to track the daily trends of glucose levels. Therefore, CGM does not replace the fingerpricking test. For all commercially available CGM devices, the sensors must be re-calibrated by fingerpricking testing on the average about 4 times per day in order to maintain measurement accuracy.

1.2.3. Transdermal Sensors

Transdermal sensors feature technologies which measure glucose molecules extracted from the ISF across the skin barrier by applying physical energy [Oliver et al., 2009]. For example, the GlucoWatch (Animas, West Chester, PA, USA) was the first FDA-approved transdermal glucose sensor, which utilized reverse iontophoresis, a technique that applies low electric current across the skin between two electrodes to excite glucose molecules to pass across the dermis layer faster than due to passive permeability. Other types of physical energizers, such as low-frequency ultrasound and skin suction have been applied as well [Kost et al., 2000]. The main factors causing measurement inaccuracies are parasitic but electrically charged molecules on the skin surface due to concurrent phenomena such as sweating, temperature fluctuations, electrostatic noise sources, etc. [Hoeks et al., 2011]. Furthermore, the electrical current has been reported causing irritations by many users. As a result, the GlucoWatch was eventually withdrawn from the market in 2008.

1.3. Innovations in Glucose Monitoring and Management

1.3.1. Non-invasive Methods

Non-invasive glucose measurements refer to technologies that measure glucose levels in ISF without causing any tissue damage. Optical sensors use light of variable wavelengths to detect glucose and utilize different properties of light to interact with glucose in a concentration-dependent manner, which include scattering, thermal infrared, fluorescence, Raman, mid-infrared and near infrared spectroscopy, etc. [Oliver et al., 2009]. Good correlations were found in the measurements. However, significant differences remain with respect to comparative laboratory blood glucose measurements, due to the inter-subject variability in skin components. Presently, none of the optical sensors have become clinically accepted and approved by the FDA, based on inferior precision and reliability as compared to the standard fingerpricking measurements. Researchers have also tried to find the correlation between blood glucose level with glucose levels in other biological fluids which are easily accessible, such as tears [Yao et al, 2011] and saliva [Agrawal et al, 2013]. While these approaches make measurements easy, they are suffering from the interference caused by other biological molecules in the sample, which are dynamically changing and environment-specific [Smith, 2013].

1.3.2. Islet Cell Transplantation

It has been hypothesized that upon the transplantation of insulin-producing islet cells from a donor pancreas into another person [Piemonti & Pileggi, 2013], the donor’s islet cells can start producing insulin to regulate blood glucose levels in the recipient's body. The concept of Islet Cell Transplantation is not new, but still remains an experimental treatment for T1DM. Two important limitations preclude its widespread application: (a) donor's islet cell rejection by the recipient's immune system is unavoidable and the transplant recipient must remain on immunosuppressant drugs [Lakey et al., 2003]; and (b) the limited supply of islet cells available for transplantation. As a result, only 471 T1DM patients were recorded receiving islet cell transplantations worldwide in the five-year period from 1999 to 2004 [Shapiro, et al., 2005]. Replacing human islet cells for transplantations with piglet islet cells only aggravated immunological rejections [Elliott, 2011].

1.3.3. Stem Cells Transplantation

Stem cell therapy for diabetes hypothesizes that patient's own stem cells can differentiate themselves into a set of islet cells which subsequently can be transplanted into the patient to produce insulin [Pagliuca, et al., 2014]. In theory, stem cell therapy can address the two problems of the islet cell transplantation approach, because there is no immunological reaction to the patients' own cells and there are plenty of stem cells suitable for the therapy. The main challenge to this new approach is that some stem cells differentiate into tumors after transplantation. This option will remain an experimental treatment until this problem is fully addressed [El-Badri & Ghoneim, 2013].

1.4. The Concept of the Electronic Mosquito

There is an obvious need for minimally-invasive blood sampling devices which would be able to automatically obtain and analyze a series of static whole blood samples over an extended period of time with minimal pain and limited user intervention. E-Mosquito is a nature-inspired concept which is implemented with advanced MEMS technologies. Its first generation was introduced about 10 years ago [Gattiker et al., 2005]. The original idea was to implement a skin-patch device which included a set of single-use cells. Figure 1 shows the detailed mechanism of an e-Mosquito cell. The device can be tightly adhered onto the skin and contains a precisely controlled actuator, which delivers a sufficient force to drive a microneedle to penetrate under the skin. A whole blood sample can be extracted as a result of the natural blood pressure gradient in the penetrated capillary vessel. A glucose sensor measures the glucose level of the extracted miniscule static whole blood sample and the signal is then sent by an RF transmitter for post-processing and display after Analog-to-Digital conversion. Each e-Mosquito device carries multiple cells of single-use needle-sensor assemblies which are to be replaced daily.

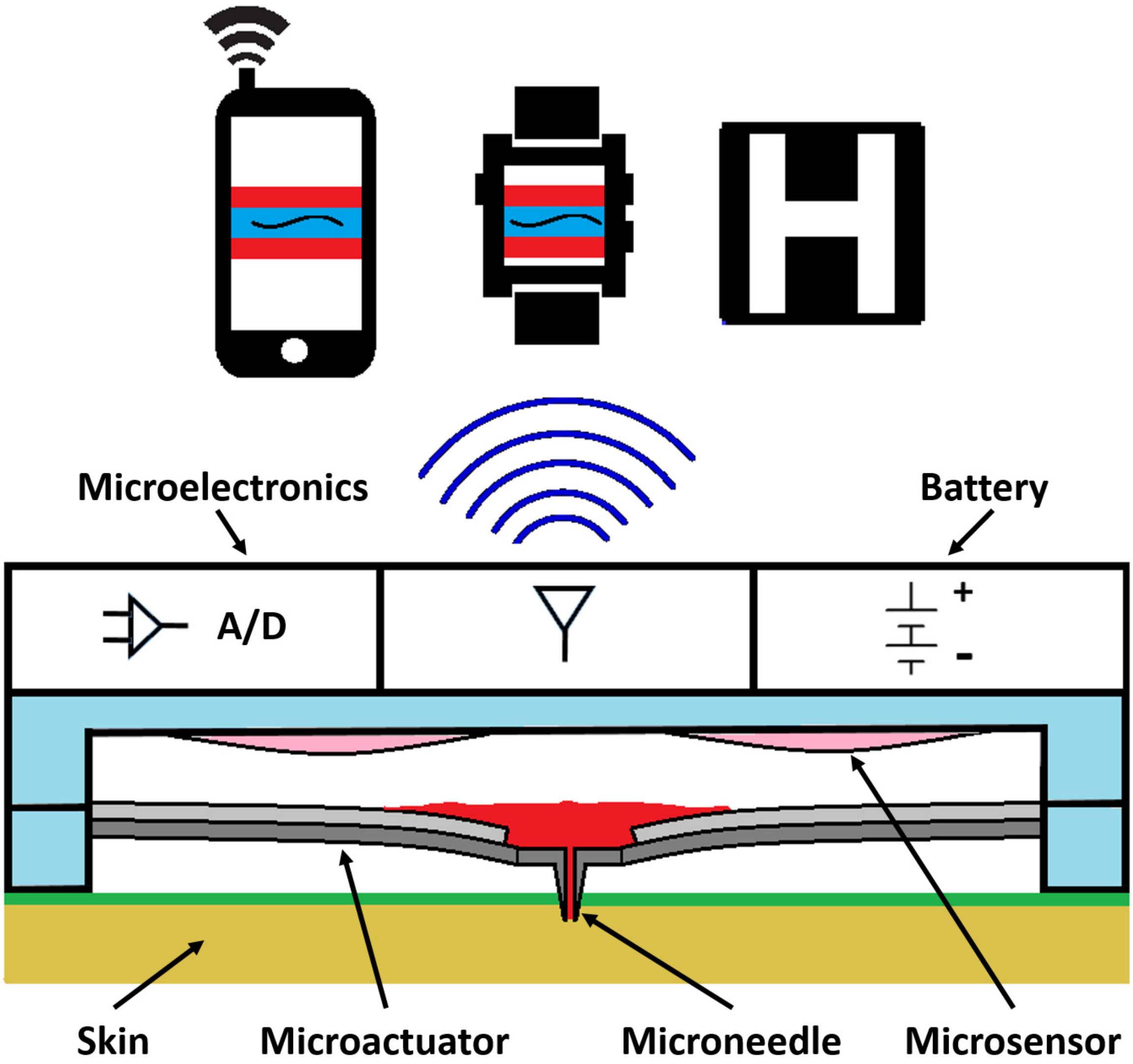


Figure 1. A single cell of the first-generation e-Mosquito prototype

In its macro-size first-generation prototype, the e-Mosquito utilized a pair of piezoelectric actuators which exerted a force of nearly 100 gf with a maximum stroke of 1.25mm [Gattiker, 2006]. An upgraded version of this system further incorporated an impedance sensor detecting the presence of a blood sample to form a closed-loop control of the actuator. It successfully extracted a blood sample of 10μl in a chicken model test [Thomas, 2009]. However, several design problems hampered its commercialization feasibility. First of all, the piezoelectric actuators required a high driving voltage of up to 60V. To achieve this voltage, additional electronics had to be integrated into the device, which greatly increased its thickness. In addition, the actuation design was not reusable due to sanitization requirements. The expensive piezoelectric actuators had to be disposed of after use, which made the device unaffordable.

The main barrier for acceptable clinical use of the first-generation e-Mosquito device was the lack of a compact, yet effective actuator to painlessly and regularly withdraw blood samples from subcutaneous capillaries. With space and energy constraints for wearable devices, it is not easy to insert a needle about 1mm below the skin where capillary vessels are abundant.

1.5. Aims of the Present Paper

This article aims at proposing a shape memory alloy (SMA) in-plane microactuator to be integrated with a miniaturized blood glucose sensing circuit and control unit in a cuff-based blood glucose monitoring device. A precise instrument is implemented to quantitatively verify that the performance of the actuator meets the requirements for capillary-reaching skin penetration. The skin penetration capability was also studied in a pilot in-vivo human experiment.

2. Methods

2.1. Design Criteria for A Microactuator for Blood Extraction

The microactuator has to meet a minimum of four pivotal design constraints to successfully fulfill its purpose as the major building block of a second-generation e-Mosquito device.

First of all, its thickness has to be restricted within millimeters in order to save space for other components of the e-Mosquito device. Therefore, the aim is to make the thickness of our microactuator to be around 6mm.

Second, the maximum penetration depth has to be big enough to reach the depth level where capillaries are abundant (0.8~1.5mm, [Hendriks et al., 2000]).

Third, the penetration force exerted on the microneedle must be sufficient to pierce through the force-resistive layers of the skin. An intensive literature review concluded that the minimum lancing force to penetrate the skin surface is approximately 30gf [Tsuchiya et al., 2005].

Fourth, the microactuator has to be easily integrated with the actuator control, glucose measuring and other building blocks of the e-Mosquito device.

There are several other secondary requirements for the actuator design, including low manufacturing cost, easy assembly procedure, biocompatibility, etc.

2.2. Cuff-based e-Mosquito Blood Glucose Monitoring System

2.2.1. Design Overview

Figure 2 shows a computer-aided design (CAD) drawing of the proposed cuff-based, two-cell e-Mosquito device. The substrate of the cuff is made of silicon rubber, which is the common material for wrist watch straps. The cuff is designed to be worn on the users’ arms or wrists. This new design features a wire-like SMA microactuator which replaces the piezoelectric actuator from the first-generation e-Mosquito device. Once heated above its transition temperature (usually by supplying an electrical current), the crystal structure of the SMA transforms from one state to another, which results in a macroscopic shrinkage by up to 8% of its original wire length. The energy released during the transformation exerts a longitudinal contraction force. Because of its shape flexibility, the SMA wire is wrapped circumferentially around the cuff surface, with its two ends connected to a power supply and electronic control circuitry on the opposite side of the cuff. This design saves space for the integrated glucose measurement circuit and reduces the overall thickness of the cuff to 6mm, as depicted in Figure 2 (right).

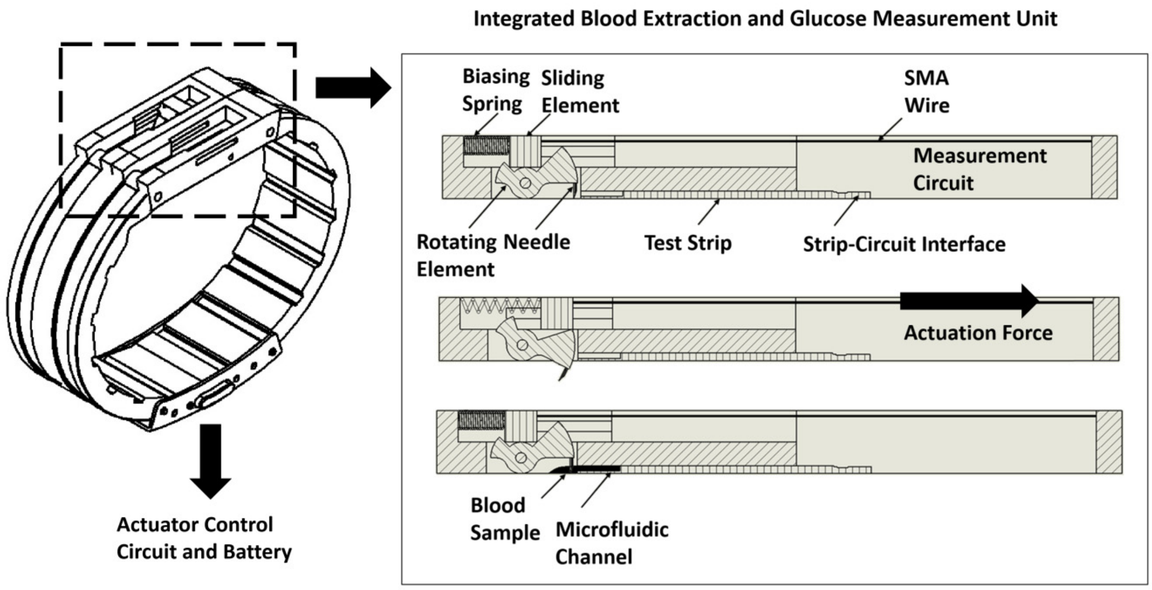


Figure 2. A conceptual prototype of the cuff-based e-Mosquito device (left); the device features a blood extraction unit utilizing a Shape Memory Alloy micro actuator (right).

2.2.2. In-plane Needle Penetration Mechanism – A Closer Look

Figure 2 depicts a simple in-plane needle-penetration mechanism, which consists of four components: a sliding element, a rotating element, a biasing spring and a needle in a planar alignment. The in-plane SMA wire is mounted on one side of the sliding element, whereas the biasing element is mounted on the opposite side of it. The sliding element is engaged with the rotating element so that when the sliding element slides under the actuation force of the contracting SMA wire, it pushes the rotating element to turn around an axis. A needle mounted on the rotating element moves off the plane and pierces into the skin at a pre-determined angle. When the shape memory effect of the SMA fiber is removed, the biasing element retracts the sliding element, which pushes the rotating element and the attached lancet back to its original position on the plane. Upon successful needle penetration reaching a capillary vessel, a small volume of blood emerges at the surface of the skin after the retraction of the needle. By capillary force, the blood flows into the microfluidic channel of a standard test strip positioned in the vicinity of the needle. The integrated glucose measurement unit then reads the glucose level in this blood sample using a standard potentiostat circuit [Wang, 2008]. The cells can be directly disposed once they are used up.

2.3. Fabrication and Assembly

The fabrication cost was taken into consideration during the prototyping process. Instead of micro-machining the microactuator components, compression molding was chosen as the prototyping method. The mold was made of stainless steel by a micro-machining system (OM-2, HAAS Automation, Oxnard, CA, USA) with a positioning accuracy of ±0.5μm and a resolution of 0.1 μm. The actuator body, sliding element and rotating element were all designed to be of equal thicknesses. Therefore, their molds were machined on the same stainless steel sheet (Figure 3).

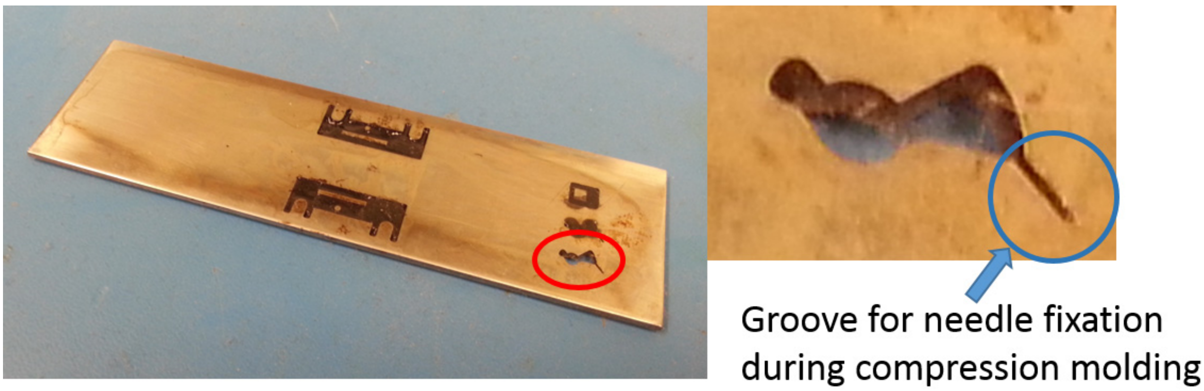


Figure 3. The mold of the actuator body, the sliding element and the rotating element (left); an enlarged photo of the circled area features a groove for needle fixation during compression molding (right).

A thermoplastic material, acrylonitrile butadiene styrene (ABS, HI-121H, Chimei, Qingdao, China), was chosen as the charge material (i.e. the plastic material that fills the mold cavity) of the compression molding because of its biocompatibility, high stiffness, good heat resistance and relatively low cost. The compression molding was performed on a manually operated four-post compression molding machine (Model 4122, Carver, Wabash, IN, USA) in approximately 30 minutes for each cycle. A 0.4mm-wide groove was machined at the front end of the rotating element for affixing a standard 33-gauge microneedle. Due to the small width of this groove, the ABS charge did not flow into it during the compression molding process. Therefore, the needle tip remained exposed and sharp after the molding.

Seven sets of the plastic components were manufactured. They were assembled with the SMA wire (BMF150, Toki Group, Tokyo, Japan) and the biasing spring (EI007A01M, Lee Spring, Brooklyn, NY, USA, Spring rate: 17gf/mm) into seven actuator cells. Two actuator cells were used for the characterization (See Section 2.4.) and five others for the pilot human study (See Section 2.5.).

2.4. Characterization of the Microactuator by A Mechanical Test Station

In most applications, SMA microactuators are activated by a voltage supply. In the present application, the voltage supply was set at a constant 3V as suggested by the SMA actuator manufacturer. However, the voltage-dependent displacement and force curves of SMA actuators have been reported to be strongly nonlinear and to exhibit hysteresis [Lee & Lee, 2000]. Furthermore, heating and cooling processes during SMA wire fabrication are usually not precisely controlled [Senthilkumar et al., 2011]. This can result in a variation in the mechanical properties of individual SMA wires from different production batches. Therefore, mechanical testing is still the most accurate way to characterize SMA actuators. We adapted a mechanical test instrument design [Chikkamaranahalli, et al., 2005] for SMA actuators, with a particular emphasis on the biasing spring. As shown in Figure 4, one horizontal load cell (XLUS88, Tecsis, Worthington, OH, USA) and a linear variable differential transformer (LVDT) sensor (MHR250, MSI, Hampton, VA, USA) monitored the in-plane actuation force and displacement of the needle.

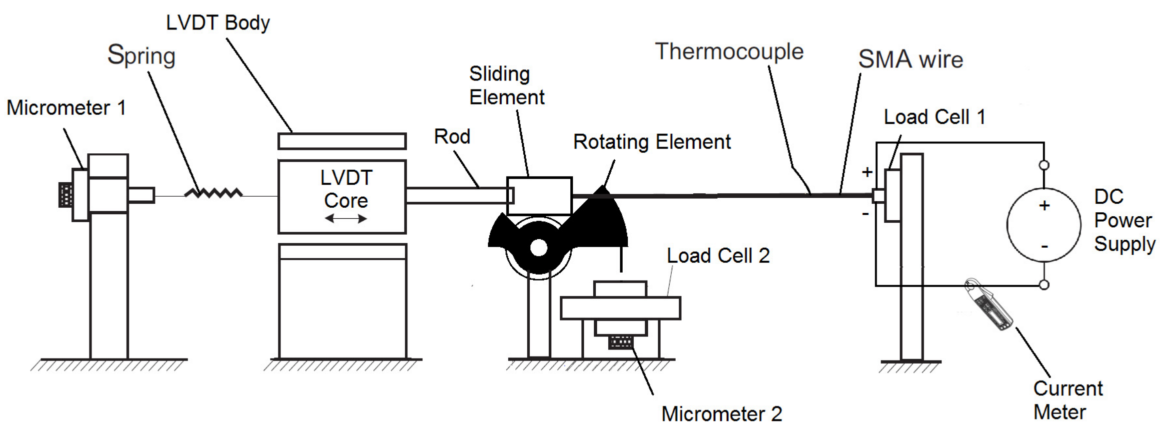


Figure 4. The mechanical test station for SMA microactuator characterization

A vertical load cell was added to the test station to keep track of the vertical force exerted by the actuator at the tip of the needle. A thermocouple (SMCJ-T, Omega, Stamford, CT, USA) and a current meter (CC-650, Hantek, Qingdao, China) provided peripheral measurements to detect whether the SMA actuator was functioning properly. A horizontal micrometer (261L, Starrett, Athol, MA, USA) was used to adjust the position of the microactuator components before each measurement. By adjusting the depth of the vertical load cell using the vertical micrometer, the penetration force of the needle at various depths was obtained.

Figure 5 shows the experimental setup of the synchronous data acquisition and control block. The central data transmission and acquisition module was implemented by a data acquisition (DAQ) card (NI-6024E, National Instruments, TX, USA) which includes a built-in timer that accurately synchronized the actuator control and monitoring processes. The input and output functions of the DAQ cards were controlled by industry-standard software (CVI, National Instruments, TX, USA). Customized analog amplification and conditioning electronic circuits for each sensor in the mechanical test station were developed for optimally utilizing the analog range of the input channels of the DAQ card. The sampling rate was set to 1 kHz for all sensors. Sensor data were stored for offline statistical analysis.

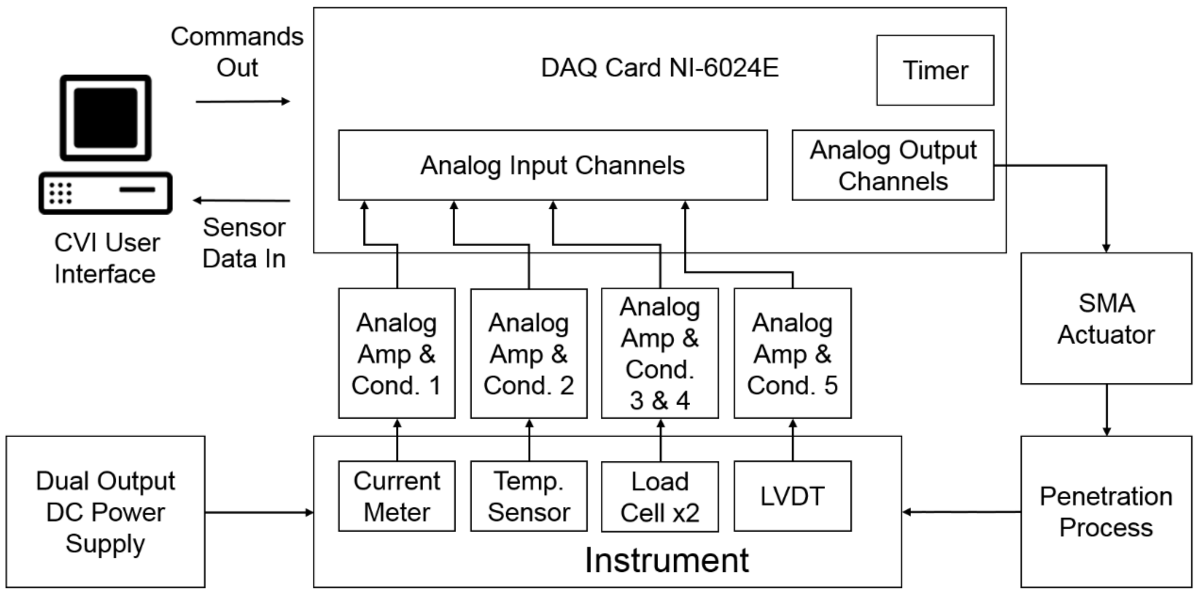


Figure 5. Block Diagram of the Control and Monitoring System

2.5. Blood Extraction Testing

One actuator cell was mounted on silicon rubber cuff (Zoro, Burnaby, BC, Canada) using miniaturized screw-and-nut mechanisms (Figure 6). The SMA wire was wrapped around the cuff and connected to a 3V DC supply via an electrical switch, which turned on and off the actuation circuit during the test. One subject (28-year old male) was recruited for the pilot blood extraction testing and signed an informed consent form. The subject was instructed to wear the cuff on his forearm and to adjust the tightness of the cuff so that the bottom surface of the actuator was in full contact with the skin surface while the cuff tightness remained at a comfortable level for him. The electrical switch was closed for 10 seconds to activate the SMA actuator during this period of time. Subsequently, the subject was instructed to remove the cuff and observe the penetration site to check if a transcutaneous bite was present and blood emerged from the bite. This test was repeated 5 times on the same subject at different locations on his left forearm.

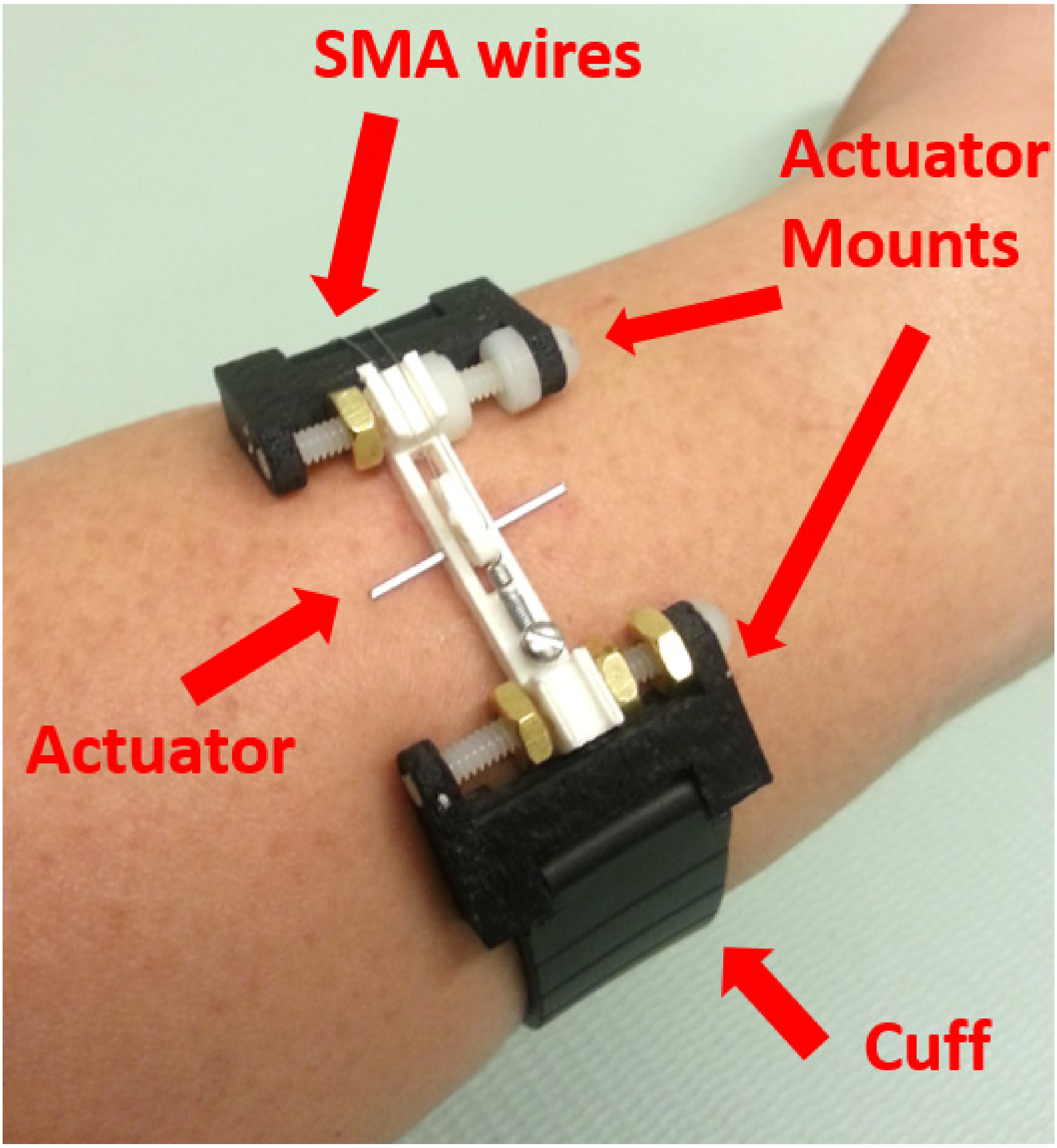


Figure 6. A functional prototype of the e-Mosquito actuator worn by the human subject on his forearm

3. Results

3.1. Prototype Implementation

Seven sets of the e-Mosquito actuator components were successfully fabricated using compression molding. The cells were assembled easily utilizing the miniaturized screw-and-nut mechanisms. Table 1 shows a breakdown of the material cost per SMA microactuator. Currently, the total cost per actuator is $3.60. This number can be easily reduced to less than $2.00 when mass-produced.

Table 1. Material cost breakdown per SMA microactuator

	Name	Unit Price ($/Unit)	No. of Units	Price
1	SMA wire	$20.00/meter	0.1 meter	$2.00
2	Biasing spring	$91.30/ 100 pcs	1 piece	$0.91
3	ABS molding granules	$2.5 / kg	0.003 kg	<$0.01
4	Gauge-33 hypodermal needle	$2.99 / 10 pcs	1 piece	$0.30
5	Mounting threaded rod and nuts	$0.19 / pair	2 pair	$0.38
	Total			**$3.60**

3.2. Characterization

We aimed to quantify two important aspects of the SMA microactuator: (1) Maximum displacement depth; and (2) Penetration force along the path of the needle movement. These two characterization tasks were accomplished by two sets of experiments using the mechanical test station.

3.2.1. Maximum Displacement Depth

The actuator was powered down and the needle was reset to its original position. Subsequently, the vertical load cell was moved 2mm below the needle position, which was used as the starting depth. At this depth, the actuator was activated 5 times. During each time, the load cell output was observed. If there was a change in the output, a counter increased by one, which was recorded into an Excel datasheet. Then, the depth of the load cell was increased by 0.025mm and the steps above were repeated at the new depth. The measurements eventually stopped at a depth of 2.5mm below the needle position. The entire test was repeated on a second SMA microactuator. Figure 7 shows a plot of the number of successful needle displacements reaching various depths recorded from the two actuators. It can be observed that the maximum penetration depth achieved by both actuators reached 2.275mm. The minimum repeatable displacement depth was not less than 2.175 mm, which was deeper than the targeted goal (0.8~1.5mm).

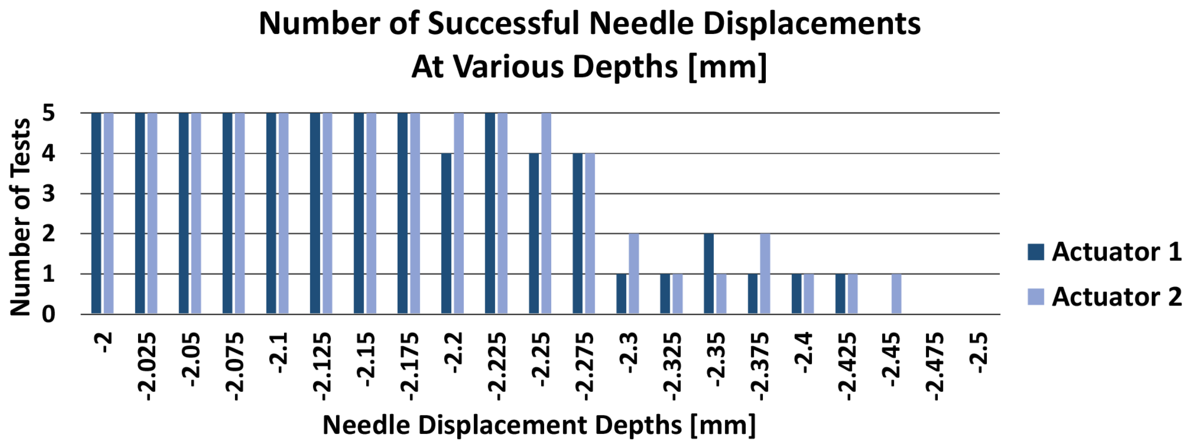


Figure 7. Maximum penetration depths of two SMA actuators (5 separate tests)

3.2.2. Penetration Force

Due to the biasing force of the spring and the change in the penetration angle during the rotation, the vertical penetration force changes at various depths below the skin surface. Therefore, the maximum exerted force of the SMA actuator alone cannot provide sufficient data to fully characterize it. In particular, force resistances of different layers of the skin vary. The top skin layer (the stratum corneum) is the most force-resistive layer compared to the layers below it [Hendriks et al., 2000]. Therefore, it is necessary to determine the penetration forces at various depths.

As described previously (see Section 3.1.), the depth of the vertical load cell can be adjusted by the micrometer mounted on it, which allows monitoring the penetration force at different depths. At the beginning of the expriment, the load cell surface was moved up to the same level as the needle's original position. Then, the SMA acutator was triggered twice. Each time, the actuation lasted 10 seconds and the maximum penetration force during these time periods was recorded into an Excel spreadsheet. Subsequently, the load cell depth was increased by 0.025 mm and the two maximum penetration forces observed at that depth were recorded again. This entire process was repeated from zero depth to the maximum displacement depth that was determined in Section 3.1. (

Figure **8**).

It can be observed from

Figure **8** that the penetration force gradually reduced as the displacement grew deeper, despite its volatile trend at shallow depths. The unstable force-depth trend at shallow depths could be explained by the hysteretic nature of the SMA microscopic transformation at low strains [Kohl, 2004]. It is possible that rapid local temperature rise leads to an explosive contraction which peaked during the recording time. The force-depth trend tended to get stabilized at deeper depths when that hysteretic period was over. The gradual reduction of the penetration force along the displacement was due to the linearly increasing biasing force by the gradually extending spring. Another reason might had been related to the increasing angle between the penetration force vector and the vertical direction as the needle moved deeper. This stems from the fact that the force measured by the load cell represented only the Z-axis component of the penetration force. Therfore, the actual penetration force at these deeper depths could had been higher than what was measured by the load cell.

Figure **8** shows that the forces were generally larger than the minimum force required for penetrating the skin [Tsuchiya et al., 2005]. The penetration force was stronger at the beginning of the penetration and consistently dropped along the penetration. This is advantageous to this particular application because the largest force is required at the top skin layer.

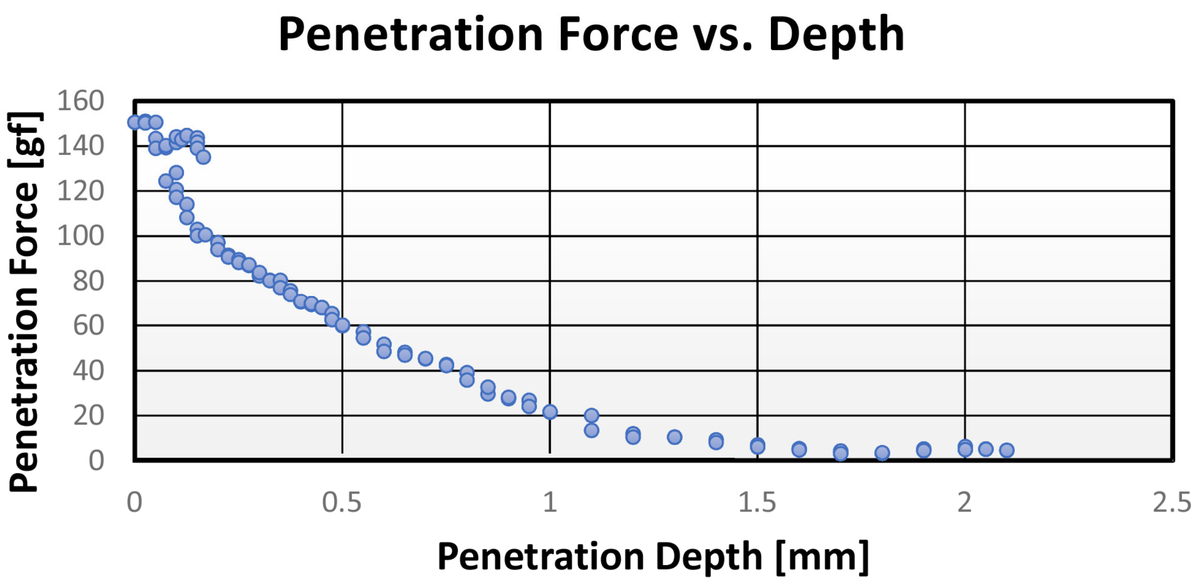


Figure 8. Penetration forces monitored by the load cell at various penetration depths

3.3. Blood Extraction Testing

The cuff-based device was removed from the subject immediately after the SMA actuation. A subsequent visual observation of the penetration site was performed for one minute after the SMA actuation. In two out of the 5 tests, a drop of blood autonomously emerged at the skin surface during the one minute. In the other three tests, a drop of blood was formed at the skin surface by mildly squeezing the skin around the penetration site. Therefore, it can be concluded that the SMA actuator successfully penetrated the skin in all of the 5 tests based on these visual observations.

Figure **9** shows a drop of blood that emerged at the skin surface after one test.

Figure 9. The SMA actuator at the activated state (left); a drop of blood (~1mm in diameter) was observed emerging at the skin surface after the device was removed (right).

4. Discussion and Future Work

An SMA microactuator was proposed for the e-Mosquito, an automatic blood extraction device aiming to provide minimally invasive glucose monitoring for diabetic patients. The actuator was prototyped and preliminarily tested in laboratory setting. There are several important issues which have to be addressed in the future. Repeatability tests on the penetration forces and depths of the proposed SMA microactuator are needed in order to statistically determine whether they meet the minimum requirements for skin penetration and blood drop extraction without any additional mechanical manipulations in humans. Considering the hysteretic nature of the SMA microscopic transformation, more experiments are necessary to better characterize this microactuator.

The feasibility of this innovative SMA actuator design for automatic blood sampling has been demonstrated in a pilot in-vivo human testing. The SMA actuators exerted sufficient force to penetrate the skin barrier and reached the capillary vessel during these 5 tests. However, the blood automatically emerged on the skin surface only in 2 out of the 5 tests. Even in these two tests, the volume of the blood samples that were observed at the skin surface seemed not sufficient for an accurate test strip testing. The small blood volume could be related to the small diameter of the 33-gauge microneedle, which was demonstrated before [Fruhstorfer et al, 1999]. In the future, microneedles with bigger diameters can be tested using the same SMA actuator. Quantitative methods have to be implemented to determine the blood volume.

Conclusion

This article proposed a novel cuff-based microactuator for automatic blood extraction using SMA technology. It aimed at providing an actuator solution for the e-Mosquito device, which could significantly enhance the quality of life for diabetic patients as well as improve their disease management. Fabrication and assembly of the microactuator were discussed. The microactuator prototype underwent some mechanical tests in laboratory environment to characterize its penetration force and displacement depth. The test results showed that this actuator met the minimum force and depth requirements for skin penetration. The feasibility of this innovative SMA actuator design for automatic blood sampling has been demonstrated in a pilot in-vivo human testing, in which the SMA actuators successfully penetrated the skin barrier and reached capillary vessels. Optimization of the needle assembly combined and quantification of the blood extraction capabilities of the actuator are needed.

Acknowledgements

This study was supported in part by the Natural Sciences and Engineering Research Council of Canada and the Innovates Centre of Research Excellence (iCORE), Alberta Innovates Technology Futures.

Bibliography

[ADA, 2012a] American Diabetes Association, “Standards of Medical Care in Diabetes – 2012”, Diabetes Cares, vol. 35, no. 2, pp. S11-63, 2012.

[ADA, 2012b] American Diabetes Association, “Diagnosis and Classification of Diabetes Mellitus”, Diabetes Care, vol. 35, no. S1, pp. S64-71, 2012.

[Agrawal et al, 2013] R.P. Agrawal, N. Sharma, M.S. Rathore, V.B. Gupta, S. Jain, V. Agarwal, S. Goyal, “Noninvasive Method for Glucose Level Estimation by Saliva”, Diabetes & Metabolism, vol. 4, no. 5, pp. 1-5, 2013.

[Bindra et al, 1991] D. S. Bindra, Y. Zhang, G. S. Wilson, R. Sternberg, D. R. Thevenot, D. Moatti and G. Reach, “Design and in vitro studies of a needle-type glucose sensor for subcutaneous monitoring.”, Anal. Chem., vol. 63, no. 17, pp. 1692-1696, 1991.

[Fruhstorfer et al, 1999] H. Fruhstorfer, G. Schmelzeisen-Redeker and T. Weiss, “Capillary blood sampling: relation between lancet diameter, lancing pain and blood volume”, European Journal of Pain, vol. 3, no. 3, pp. 283-286, 1999.

[Gattiker et al, 2005] G. Gattiker, K. Kaler and M. Mintchev, “Electronic Mosquito: designing a semi-invasive Microsystem for blood sampling, analysis and drug delivery applications”, Microsyst Technol, vol. 12, no. 1, pp. 44-51, 2005.

[Gattiker, 2006] G. Gattiker, PhD Thesis: Designing a BioMEMS-based Blood Sampler, Calgary: University of Calgary, 2006.

[Hendriks et al, 2000] F. M. Hendriks, D. Brokken, C. W. J. Oomens, F. P. T. Baaijens and J. B. A. Horsten, “Mechanical Properties of Different Layers of Human Skin”, Philips Research Laboratories, Eindhoven, 2000.

[Hoeks et al, 2011] L. Hoeks, W. Greven and H. d. Valk, “Real-time continuous glucose monitoring system for treatment of diabetes: a systematic review”, Diabet Med, vol. 28, no. 2, pp. 386-94, 2011.

[Kohl, 2004] M. Kohl, Shape Memory Microactuators, Berlin: Springer, 2004.

[Kost et al, 2000] J. Kost, S. Mitragotri, R. A. Gabbay, M. Pishko and R. Langer, “Transdermal monitoring of glucose and other analytes using ultrasound”, Nat Med, vol. 6, no. 3, pp. 347-350, 2000.

[Lakey et al, 2003] J. Lakey, P. Burridge and A. Shapiro, “Technical aspects of islet preparation and transplantation”, Transpl Int , vol. 16, no. 9, pp. 613-632, 2003.

[Lee and Lee, 2000] H. Lee and J. Lee, “Evaluation of the characteristics of a shape memory alloy spring actuator”, Smart Materials and Structures, vol. 9, no. 6, pp. 817-823, 2000.

[Oliver et al, 2009] N. S. Oliver, C. Toumazou, A. E. G. Cass and D. G. Johnston, “Glucose sensors: a review of current and emerging technology”, Diabetic Medicine, vol. 26, no. 3, pp. 197-210, 2009.

[Pagliuca et al, 2014] F. W. Pagliuca, J. R. Millman, M. Gürtler, M. Segel, A. Van Dervort, J. Ryu, Q. P. Peterson, D. Greiner and D. A. Melton, “Generation of functional human pancreatic β cells in vitro.”, Cell, vol. 159, no. 2, pp. 428-39, 2014.

[Simmons, 2008] B. Simmons, "Micromolding (Injection and Compression Molding)," in Encyclopedia of Microfluidics and Nanofluidics, New York, Springer US, 2008, pp. 1267-1274.

[Smith, 2013] J. L. Smith, “The Pursuit of Noninvasive Glucose: Hunting the Deceitful Turkey”, Internet: http://www.mendosa.com/The%20Pursuit%20of%20Noninvasive%20Glucose%203rd%20Edition.pdf, [Apr 1, 2015]

[Thomas, 2009] G. Thomas, MSc Thesis: Electronic Mosquito: A Feedback-Controlled Semi-Invasive Microsystem for Glucose Monitoring, Calgary: University of Calgary, 2009.

[Tsuchiya et al, 2005] K. Tsuchiya, N. Nakanishi, Y. Uetsuji and E. Nakamachi, "Development of Blood Extraction System for Health Monitoring System", Biomedical Microdevices, vol. 7, no. 4, pp. 347-353, 2005.

[Yao et al, 2011] H. Yao, A.J. Shum, M. Cowan, I, Lahdesmaki, B.A. Parviz, "A contact lens with embedded sensor for monitoring tear glucose level", Biosensors and Bioelectronics, vol. 26, no. 7, pp 3290-3296, 2011.

Authors' Information

Gang Wang *is a PhD candidate in the Biomedical Engineering Graduate Program, University of Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada; email: gawang@ucalgary.ca. Major Fields of Scientific Research: Biomedical Instrumentation and Imaging.*

Michael D. Poscente *is a Master student in the Biomedical Engineering Graduate Program, University of Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada; email: mdposcen@ucalgary.ca. Major Fields of Scientific Research: Biomedical Instrumentation.*

Simon S. Park *is a professor in the Department of Mechanical and Manufacturing Engineering, University of Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada; e-mail: sipark@ucalgary.ca. Major Fields of Scientific Research: Nano-MEMS, Automation, Manufacturing engineering, Design, Vibration.*

Orly Yadid-Pecht *is a professor in the Department of Electrical and Computer Engineering, University of Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada; e-mail: Orly.Yadid-Pecht@ucalgary.ca. Major Fields of Scientific Research: Image Sensors, Photonics, Micro-nano sensory systems, Biomedical engineering.*

Martin P. Mintchev *is a professor the Department of Electrical and Computer Engineering, and the Faculty of Medicine, University of Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada; e-mail: mintchev@ucalgary.ca. Major Fields of Scientific Research: Biomedical engineering, Oilfield instrumentation.*

INTELLIGENT TRADING SYSTEMS

Luis F. de Mingo, Nuria Gómez Blas, Alberto Arteta

Abstract: *During last 50 years, the markets have been object of study. Technical and fundamental indicators have been used to try to predict the behavior of the market and then execute buying or selling orders. Neural networks are currently being used with good results although they can be useless after a period of time. This paper proposes an algorithm that combines bioinspired techniques to maximize the hits in the prediction rates. The proposal shown in this paper relies in an ANN to achieve these goals. The differential factors of this approach are the election of the ANN structure with grammatical swarm and the training process through the use of HydroPSO. Also a grammatical swarm algorithm is used to generate trading rules, this method shows better results than the first approach. This combination of techniques provides an automatic way to define the most suitable bioinspired model for the instrument in our analysis.*

Keywords: *Trading strategies, Technical indicators, Grammatical Swarm, Neural networks.*

ACM Classification Keywords: *F.1.1 Theory of Computation - Models of Computation, I.2.6 Artificial Intelligence.*

Introduction

Natural sciences, and especially biology, represented a rich source of modeling paradigms. Well-defined areas of artificial intelligence (genetic algorithms, neural networks), mathematics, and theoretical computer science (L systems, DNA computing) are massively influenced by the behaviour of various biological entities and phenomena. In the last decades, new emerging fields of so-called natural computing identify new (unconventional) computational paradigms in different forms [Bonabeau et al., 1999]. There are attempts to define new mathematical and theoretical models inspired by nature. Moreover, computational paradigms suggested by biochemical phenomena [Kennedy et al., 2001] are object of study.

Genetic algorithms (GAs) are problem solving methods (or heuristics) that mimic the process of natural evolution. These algorithms utilize the concepts of natural selection to determine the best solution for a problem. As a result, GA are commonly used as optimizers that adjust parameters to minimize or maximize some feedback measure, which can then be used independently or in the construction of a neural network (ANN). Their use can be helpful when optimizing the chances for a price to rise (or decrease) [Prasad, 2004] and [Lin et al., 2004].

Stock market forecasting has been implemented using different approaches in the literature. Multilayer perceptrons [de Oliveira et al., 2014], self-organizing maps [Sarlin, 2014], radial basis functions neural networks [Zhang and Liao, 2014], support vector machines [Han and Rung-Ching, 2007] or even social learning [Xiaoac et al., 2014]. Most of them are based on a data training set and the model tries to approximate/forecast the output. In this paper we proposed a model that generates a rule set that can be used in an automatic trading system to place market orders (buy/sell).

Basic Stock Market Indicators

In statistics, a moving average (rolling average or running average) is a calculation to analyze data points by creating a series of averages of different subsets of the full data set. It is also called a moving mean (MM) [Booth et al., 2006] or rolling mean and is a type of finite impulse response filter. Variations include: simple, and cumulative, or weighted forms (described below).

Given a series of numbers and a fixed subset size, the first element of the moving average is obtained by taking the average of the initial fixed subset of the number series. Then the subset is modified by "shifting forward"; that is, excluding the first number of the series and including the next number following the original subset in the series. This creates a new subset of numbers, which is averaged. This process is repeated over the entire data series. The plot line connecting all the (fixed) averages is the moving average. A moving average is a set of numbers, each of which is the average of the corresponding subset of a larger set of datum points. A moving average may also use unequal weights for each datum value in the subset to emphasize particular values in the subset.

A moving average is commonly used with time series data to smooth out short-term fluctuations and highlight longer-term trends or cycles. The threshold between short-term and long-term depends on the application, and the parameters of the moving average will be set accordingly. For example, it is often used in technical analysis of financial data, like stock prices, returns or trading volumes. It is also used in economics to examine gross domestic product, employment or other macroeconomic time series. Mathematically, a moving average is a type of convolution and so it can be viewed as an example of a low-pass filter used in signal processing. When used with non-time series data, a moving average filters higher frequency components without any specific connection to time, although typically some kind of ordering is implied. Viewed simplistically it can be regarded as smoothing the data.

HydroPSO

HydroPSO is an enhanced version of the canonical Particle Swarm Optimization (PSO) technique which was implemented and developed by [Kennedy and Ebehart, 1995]. As explained, PSO is a population-based stochastic optimization technique inspired by social behaviour of bird flocking. This is part of the of evolutionary optimization techniques such as Genetic Algorithms (GA). The main problem PSO faces is the exploration within a multi-dimensional solution space due to the lack of evolution operators which can endanger the adaptability of the neural network by falling into local optima. HydroPSO, however is capable of performing high sensitive analysis just by using the Latin Hypercube One-At-a-Time (LH-OAT) method [van Griensven et al., 2006; Iman, 2008; Singh et al., 2010; Penga and Lua, 2013], the power of HydroPSO is the increase of ability to adjust better to new models.

The LH-OAT provides a constrained sampling scheme instead of random sampling according to the direct Monte Carlo simulation. In the LH-OAT, the region is uniformly divided into N non-overlapping intervals for each random variable; where N is the number of random numbers. These need to be generated for each random variable. The N non-overlapping intervals are selected to be have the same probability of occurrence. Then, N different values in the N non-overlapping intervals are randomly chosen for each random variable. This can be accomplished by generating N random numbers in the first place. These represent the percentage position of each generated value corresponding to the variable within an interval. In the bottom layer of HydroPSO, several swarms evolve in a parallel way to avoid being trapped in local optima. The learning strategy for each swarm is the well-known comprehensive learning method with a newly designed mutation operator. As the evolution processes at the bottom layer, one particle for each swarm is selected as a candidate to construct the swarm in the top layer, which evolves by the same strategy employed in the bottom layer. The local search strategy based on LH-OAT is imposed on particles in the top layer every specified number of generations

Neural Network training using HydroPSO

The working dynamic ANN can be performed just by running these two following processes: creating a neural network and finding the right weights by using the HydroPSO package. Given a neural network architecture, every weight is coded as a genotype; then the network is trained by running the Hydroparticle swarm optimization algorithm.

First, three random particles (r_1, r_2, r_3) are selected, note that they must be different from each other. The second task is to create a mutated value for each dimension j of the particle according to the evolutionary algorithm.

New particles are created according to standard PSO formulas: Velocity and new position. A basic velocity and position clamping is performed as well. Then the new values (velocity, position) are checked to determine whether they exceed the threshold defined in the algorithm. If the newly created particle is proven to be better, it will replace the old one for the next generation. Both personal best and global best particle vectors are updated as well. This whole process occurs over again until the Halt condition is reached.

In our case, the solutions of the ANN are optimized by HydroPSO algorithm. The fitness function to be optimized is :

$$\phi(i) = \frac{|x_i - y_i|}{x_i} \tag{1}$$

here, x_i is the particle that represents the real value that a market instrument gets in the moment i and y_i is the particle representing the value that our ANN returns.

Equations used in the particle swarm optimization training process are below; c_1 and c_2 are two positive constants, R_1 and R_2 are two random numbers belonging to $[0, 1]$ and w is the inertia weight. These equations define how the genotype values change along iterations; In other words this equations show how neural network weights change.

$$x_{in}(t+1) = x_{in}(t) + v_{in}(t+1) \tag{2}$$

The equations above make the network weights gets updated until the Halt condition is reached, that is to say, either the error is minimized (fitness function is closed to zero) or a maximum number of iterations is reached.

This basic example reveals that the previously defined PSO algorithm can be successfully applied in the training stage in order to solve the convergence problem when working with high dimension individuals. The XOR example with dimension 9 is a good candidate to start with combining classical neural networks with Swarm intelligence.

Grammatical Swarm

Grammatical Swarm (GS) [O'Neill and Brabazon, 2006] relates Particle Swarm algorithm to a Grammatical Evolution (GE); genotype-phenotype mapping to generate programs in an arbitrary language [O'Neill and

Brabazon, 2004]. The equations for the particle swarm algorithm are updated by adding new constraints to velocity and location dimension values, such us vmax (bounded to ± 255), and search space dimensions which are bounded to the range $[0, 255]$ (This is denoted as cmin and cmax, respectively). Note that this is a continuous swarm algorithm with real-valued particle vectors. The standard GE mapping function is adopted, with the real-values in the particle vectors being rounded up or down to the nearest integer value for the mapping process. In the current implementation of GS, fixed-length vectors are used, which implies that it is possible for a variable number of dimensions to be used during the program construction genotype-phenotype mapping process. A vector's elements (values) may be used more than once if wrapping occurs, and it is also possible that not all dimensions are used during the mapping process. (This can happen whenever a program is generated before reaching the end of the vector).

Neural Network topology using GS: first approach

Previous PSO model applied to a fixed neural network is a good training solution, however it does not define any kind or topology properties as it only obtains the best weight values. Following grammars can be used with Grammatical Swarm algorithms in order to obtain a network topology for a given problem.

This grammar can specify a feed-forward neural network topology with consecutive layers, that is to say, a classical Multilayer Perceptron.

```
<layers> ::= <layer> | <layer>, <layers>
<layer> ::= <digit>
<digit> ::= 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9
```

Next grammar is able to generate feed-forward connections not only with one consecutive layer but also with more than one consecutive layer. Such connections are defined by the `<connections>` non terminal, where the `<digit>` means the n-consecutive layer.

```
<layers> ::= <layer> | <layer>, <layers>
<layer> ::= <digit> -- <connections> --
<connections> := <digit> | <digit>, <connections>
<digit> ::= 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9
```

The whole algorithm is summarized as follows:

1. Create an initial population of genotypes.
2. For genotype i
   (a) Using genotype and grammar to obtain a neural architecture.
   (b) Compute Fitness of genotype.
      - Apply previous HydroPSO algorithm to train the genotype network.
   (c) Modified the best individual if appropiate.
3. Update velocity of genotype i.
4. Update position of genotype i.
5. If stop condition is not satisfied go to step 2.

This neural network model is a powerful one as a network topology is chosen and trained only with the input and output pattern data sets. Both tools, topology and training, are based on grammatical swarm and particle swarm optimization respectively.

Another useful approach could be to defined a grammar with the weight values of neural networks connections instead of training the topology using a PSO algorithm.

Trading Results

Technical analysis utilizes models and trading rules based on price and volume transformations, such as the relative strength index, moving averages, regressions, inter-market and intra-market price correlations, business cycles, stock market cycles or, classically, through recognition of chart patterns. Technical analysis stands for mathematical indicators (volume, prices and functions) in contrast to the fundamental analysis approach (status of the companies, government decisions, etc).

Technical analysis is widely used among traders and financial professionals and is very often used by active day traders, market makers and pit traders. In the 1960s and 1970s technicality was widely dismissed by academics; however in recent reviews, Irwin and Park reported that 56 of 95 modern studies found that Technical studies produces positive results. A more detailed explanation of the ANN achievements in the stock markets can be found in [Dase and Pawar, 2010]

Data are obtained from `tradingmotion.com`, it provides the stock price every 30 minutes from 2002 until today, and so the common trading indicators, such as RSI, MACD, BBands, etc. There are more than 50000 patterns (that is, 12 years times 16 market ticks a day (8 hours)). We have randomly taken 90% of these for training purposes and 10% for testing purposes.

Next sections will show the results obtained in our analysis by using (neural networks and grammatical swarm) when applied to the stocks market.

Neural Networks with HydroPSO learning and Grammatical Swarm architecture

This neural network has been trained to forecast the IBEX indicator. Data from 2002 to 2014, obtained every 30 minutes, are used. There are 55490 patterns. Stock market dataset is made of 176 inputs (16 trading indicators from t to $t-10$) and 2 outputs corresponding to the price change: $Price_t - Price_{t+1}$ and $Price_t - Price_{t+5}$. All data are normalized in interval $(-1, 1)$. Following list shows the indicators: *Price change, EMA, Cross average, MACD line, MACD signal, Bollinger bandwidth, Bollinger center distance, Momentum, RSI, ATR, Aroon up, Aroon down, CCI, Stochastic K, Stochastic D, EOM.*

The best performance has been achieved with a neural network (multilayer perceptron with 100 hidden neurons) trained with a HydroPSO algorithm and a grammatical swarm one, in comparison with the classical back propagation algorithm. Figures 1, 2 and tables 1, 2 show obtained results.

Trading rules by using Grammatical Swarm

In this section we display the results when using a BNF grammar to generate a evolutionary algorithm implemented in C++. At first 16 trading indicators are used; and then the ones with with a low influence in the forecasting results are discarded

Table 1: Neural network training performance

Performance results	
Main chart: IBEX 30 minute bars [27/12/2001 - 08/03/2013]	
Net P&L	25183,09 €
Gross P&L	39280,00 €
Profit factor	1,10040504
Sharpe ratio	0,215754143
Slippage per side	-2,04524309
Commission per side	5,75 €
Annual ROI	1,95%
Mathematical expectation	73,01115242
Analyzed sessions	2841
Sessions in market	771
Winning sessions	430
Winning sessions profit	275998,11 €
Winning sessions average	641,86 €
Losing sessions	341
Losing sessions profit	-250815,02 €
Losing sessions average	-735,53 €
Worst drawdown	-31216,25 €
Best session	3570,00 €
Worst session	-5140,00 €

Next list shows the trend which was detected by the grammatical swarm algorithm; After that, ADX, MACD and Bollinger Bands were used, although the only relevant indicator turned out to be the ADX. A graphical data flow is shown in figure 3.

```
if ((adx.GetADX()[0]<30) && (adx.GetADX()[1]>30)) {
  trend = false;
  high_trend = false;
}
else if ((adx.GetADX()[1]<30) && (adx.GetADX()[0]>30)) {
  trend = true;
  high_trend = false;
}
else if ((adx.GetADX()[0]<45) && (adx.GetADX()[1]>45)) {
  high_trend = false;
  trend = false;
}
else if ((adx.GetADX()[0]>45) && (adx.GetADX()[1]<45)) {
  high_trend = true;
  trend = true;
}
```

The following code shows the buy/sell orders depending of a given indicators that have been generated by the grammatical swarm.

```
if (this.GetOpenPosition() == 0) {
  if (trend && !high_trend) {
   if ( (adx.GetPlusDI()[1] < adx.GetPlusDI()[0]) &&
        (adx.GetMinusDI()[1] > adx.GetMinusDI()[0]) &&
        (adx.GetPlusDI()[0] > adx.GetADX()[0]) &&
        (adx.GetMinusDI()[0] < adx.GetADX()[0] ) &&
        (macd.GetSignalHistogram()[0] > 0)
    )
   this.Buy(OrderType.Market, 1, 0, "Entry Long");
```

Table 2: Neural network testing performance

Performance results	
Main chart: IBEX 30 minute bars [27/12/2012 - 07/03/2014]	
Net P&L	13009,68 €
Gross P&L	15110,00 €
Profit factor	1,712780504
Sharpe ratio	1,744592161
Slippage per side	-2,812612282
Commission per side	5,75 €
Annual ROI	54,58 %
Mathematical expectation	243,7096774
Analyzed sessions	305
Sessions in market	84
Winning sessions	51
Winning sessions profit	31261,70 €
Winning sessions average	612,97 €
Losing sessions	33
Losing sessions profit	-18252,01 €
Losing sessions average	-553,09 €
Worst drawdown	-2735,76 €
Best session	2400,00 €
Worst session	-1762,17 €

```
    if ((adx.GetMinusDI()[1] < adx.GetMinusDI()[0]) &&
        (adx.GetPlusDI()[1] > adx.GetPlusDI()[0]) &&
        (adx.GetMinusDI()[0] < adx.GetADX()[0]) &&
        (adx.GetPlusDI()[0] > adx.GetADX()[0]) &&
        (macd.GetSignalHistogram()[0] < 0)
      )
    this.Sell(OrderType.Market, 1, 0, "Entry Short");
  }
} else if (this.GetOpenPosition() == 1) {
  if (trend)
     this.ExitLong(OrderType.Stop,
     bbands.GetLowerBand()[0], "Exit Long");
  else if (!high_trend)
     this.ExitLong(OrderType.Limit,
     bbands.GetUpperBand()[0], "Exit Long");
}  else if (this.GetOpenPosition() == -1) {
  if (trend)
     this.ExitShort(OrderType.Stop,
     bbands.GetUpperBand()[0], "Exit Short");
  else if (!high_trend)
     this.ExitShort(OrderType.Limit,
     bbands.GetLowerBand()[0], "Exit Short");
}
```

Figures 4, 5 and tables 3, 4 show obtained results in training and testing.

The results reveal that grammatical approach behaves better than the neural network approach when forecasting IBEX indicator. The number of hits (winning sessions) is clearly higher when using grammatical swarm with HydroPSO. In this scenario a hit is considered a winning transaction that has been placed in the market with the estimated price that the ANN has returned. HydroPSO, as well as PSO, improve the performance when training the ANN which has been elected through grammatical swarm. The election of ANN through G.S. ensures the most suitable architecture is chosen in regards of desired outcomes. Furthermore, the use of HydroPSO, helps in the performance when training the ANN. Finding the best

Table 3: Grammatical swarm training performance

Performance results	
Main chart: IBEX 30 minute bars [27/12/2001 - 08/03/2013]	
Net P&L	63069,63 €
Gross P&L	79520,00 €
Profit factor	1,24649621
Sharpe ratio	0,623895837
Slippage per side	-2,069753347
Commission per side	5,75 €
Annual ROI	11,26%
Mathematical expectation	127,8456592
Analyzed sessions	2841
Sessions in market	866
Winning sessions	511
Winning sessions profit	318934,15 €
Winning sessions average	624,14 €
Losing sessions	355
Losing sessions profit	-255864,52 €
Losing sessions average	-720,75 €
Worst drawdown	-12503,25 €
Best session	4630,00 €
Worst session	-4425,75 €

topology to get the most accurate results and optimizing the training process contribute reveal themselves as a reliable combination to optimize the ANN functionality, as shown in the results.

Table 5 show the performance of four algorithms. The first one is the proposed grammatical swarm with a performance of 92.85%, next column shows the Neural Network based algorithm with a performance of 54.58%, then the Golden Cross algorithm (based on a basic stock market rule) with -12.7% and finally the Trend A1 algorithm with 15.2%. The last one is the best automatic system (among 39 systems), please check following web page `www.tradingmotion.com` (see figure 6).

There are a lot of indicators to evaluate a trading system, depending on the behaviour: (From high risk to low risk). `Annual ROI` (The total return divided by the number of years in the period, total return = net profit/loss divided by suggested capital), `Profit factor` (The profit factor is the ratio between profits and losses, and its calculated by dividing the sum of profits over the sum of losses), `Net P&L` (The total profit or loss (P/L) over the analysed period, net of commissions, slippage and license costs), `Worst drawdown` (The worst peak to valley loss of the system, as measured on an end of session basis, with the date of the low point listed), `Sharpe ratio` (Sharpe ratio measures the excess return per unit of deviation. It characterizes how well the return of an asset compensates the investor for the risk taken. When comparing two assets versus a common benchmark, the one with a higher Sharpe ratio provides better return for the same risk. Sortino is a modification of the Sharpe ratio but penalizes only those returns falling below a specified target, or required rate of return, while the Sharpe ratio penalizes both upside and downside volatility equally), etc...

By checking the mentioned parameters we conclude that the proposed grammatical swarm behaves, in general, better than many automatic system described in `tradingmotion.com`

Conclusion and further work

This paper analyzes a few optimization strategies in the natural computation area such as competitive and collaborative models. (ANN, Grammatical Swarm and HydroPSO) These have been described in order to extract some biological ideas and then apply them in computational models. Such bio-inspired models

Table 4: Grammatical swarm testing performance

Performance results	
Main chart: IBEX 30 minute bars [27/12/2012 - 07/03/2014]	
Net P&L	16597,88 €
Gross P&L	18580,00 €
Profit factor	3,063432494
Sharpe ratio	2,772734456
Slippage per side	-2,728527645
Commission per side	5,75 €
Annual ROI	92,85 %
Mathematical expectation	309,6666667
Analyzed sessions	305
Sessions in market	73
Winning sessions	52
Winning sessions profit	24641,71 €
Winning sessions average	473,88 €
Losing sessions	21
Losing sessions profit	-8043,82 €
Losing sessions average	-383,04 €
Worst drawdown	-1075,58 €
Best session	1560,42 €
Worst session	-1180,00 €

have proven to be an effective tool for solving non common problems; As a powerful application, neural networks can take advantage of grammatical swarm optimization models. This paper has introduced an optimized architectural neural network (built up through grammatical swarm techniques) trained by a HydroPSO algorithm. In particular this paper has shown that building an ANN with Grammatical Swarm and then trains it with HydroPSO provides a good theoretical outcome in terms of winning transactions in the stock market. The solution provided is based on the historical data of the last 2 years. Since 09/03/2014 the predictions rates behave extremely well, but of course it is impossible to know how the ANN will work from 2015 in advance. The results are promising and generalization is a solid option since the ANN architecture is optimized through the use of grammatical swarm. That ensures the optimal election of the ANN. When the ANN is not behaving well, the dynamic inherent nature of the proposed system, allows modifying the topology and restart the training process in case of a wrong prediction rates stream. That is why this process can generalize as it adapts when a pre-defined threshold is reached.

Particle Swarm Optimization often fails when searching the global optimal solution when the objective function has a large number of dimensions. The reason of this phenomenon is not just existence of the local optimal solutions but also the degenerative process of the particles velocities (this means that particles stay in a subregion within the search area)[Rapaic et al., 2009]. This is a sub-plane which is defined by a finite number of particle velocities. local optima problem in PSO is object of study. New proposals of modifications on the basic particle driven equation [Parsopoulos et al., 2001; Hendtlass, 2005] try to overcome this. They use a randomized method (e.g. mutation in evolutionary computations) for either to maintain particles velocities or to accelerate them. Although such improvements work well and have ability to avoid falls in the local optima, the problem of early convergence by the degeneracy of some dimensions still exists, even when local optima do not exist. Hence the PSO algorithm does not always work well for the high-dimensional function. That is the reason why this proposal counts on HydroPSO as the optimal trainer for ANNs. The obtained results suggest HydroPSO might be the right procedure [Zambrano-Bigiarini and Rojas, 2013].

HydroPSO allows the modeller to perform a standard modelling work flow including, sensitivity analysis, parameter calibration, and assessment of the calibration results, using a single piece of software. HydroPSO implements several state-of-the-art enhancements and fine-tuning options to the Particle Swarm Optimisation

Table 5: Results using different algorithms, note that the proposed grammatical swarm one obtains the best performance. Data obtained from `www.trading motion.com` in which the best IBEX automatic system is the Trend A1 IBEX 11.

Performance results	**Gram. Swarm**	Neural Net	Gold Cross	Trend A1 Ibex'11
Main chart: IBEX 30 minute bars [27/12/2012 - 07/03/2014]				
Net P&L	16597,88 €	13009,68 €	-55766 €	5149 €
Gross P&L	18580,00 €	15110,00 €	-57766 €	7021 €
Profit factor	3,063432494	1,712780504	0.76	1.12
Sharpe ratio	2,772734456	1,744592161	-2.0072	0.6420
Slippage per side	-2,728527645	-2,812612282	-2.1687	-2.2632
Commission per side 5,75 €				
Annual ROI	92,85 %	54,58 %	-12.7 %	15.2 %
Mathematical expectation	309,6666667	243,7096774	74,3984991	124.23984
Analyzed sessions 305				
Sessions in market	73	84	161	168
Winning sessions	52	51	42	80
Winning sessions profit	24641,71 €	31261,70 €	27541,14 €	57584.92 €
Winning sessions average	473,88 €	612,97 €	665.45 €	720.22 €
Losing sessions	21	33	119	88
Losing sessions profit	-8043,82 €	-18252,01 €	-83307.14 €	-52345.92 €
Losing sessions average	-383,04 €	-553,09 €	-700.06 €	-594.84 €
Worst drawdown	-1075,58 €	-2735,76 €	-64444 €	-5525 €
Best session	1560,42 €	2400,00 €	2362 €	2181 €
Worst session	-1180,00 €	-1762,17 €	-1783 €	-1513 €

(PSO) algorithm to meet specific user needs. HydroPSO easily interfaces the calibration engine to different model codes through simple ASCII files and/or R wrapper functions for exchanging information on the calibration parameters. Then, optimises a user-defined goodness-of-fit measure until a maximum number of iterations or a convergence criterion are met. Finally, advanced plotting functionalities facilitate the interpretation and assessment of the calibration results. The current HydroPSO version allows easy parallelization and works with single-objective functions, with multi-objective functionalities being the subject of ongoing development. Although the application of hydroPSO is mainly focused to hydrological models, flexibility of the package can be implemented in a wider range of models requiring some form of parameter optimisation, such as proposed algorithms in this paper.

Although the proposed ANN outputs are considered as a relevant indicator to obtain buying or selling signals, there is a lot more work to be done as the combination with other known indicators might contribute with an even stronger strategy for succeeding with market operations. Thus, this ANN could be an interesting part of a strategy that combines human trading experience, dynamic ANN output and conventional indicators.

Acknowledgements

The research was supported by the Spanish Research Agency projects TRA2010-15645 and TEC2010-21303-C04-02.

Bibliography

Bonabeau, E., Dorigo, M., and Theraulaz, G. (1999). *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press.

Booth, E., Mount, J., and Viers, J. H. (2006). Hydrologic variability of the cosumnes river floodplain. *San Francisco Estuary and Watershed Science*, 4(2).

Dase, R. K. and Pawar, D. D. (2010). Application of artificial neural network for stock market predictions: A review of literature. *International Journal of Machine Intelligence*, 2(2):14–17.

de Oliveira, F. A., Nobre, C. N., and Zárate, L. E. (2014). Applying artificial neural networks to prediction of stock price and improvement of the directional prediction index - case study of petr4, petrobras, brazil. *Expert Systems with Applications*, 40(18):7596–7606.

Han, S. and Rung-Ching, C. (2007). Using svm with financial statement analysis for prediction of stocks. *Communications of the IIMA*, 7(4):63–72.

Hendtlass, T. (2005). A particle swarm algorithm for high dimensional, multi-optima problem spaces. In *Proceedings of Swarm Intelligence Symposium*, pages 149–154.

Iman, R. L. (2008). *Latin hypercube sampling*. Wiley Online Library.

Kennedy, J. and Ebehart, R. (1995). Particle swarm optimization. In *Proceedings of IEEE International Conference on Neural Networks*, volume IV, pages 1942–1948.

Kennedy, J., Eberhart, R., and Shi, Y. (2001). *Swarm Intelligence*. Morgan Kauff-man.

Lin, L., Cao, L., Wang, J., and Zhang, C. (2004). *The Applications of Genetic Algorithms in Stock Market Data Mining Optimization*. WIT Press.

O'Neill, M. and Brabazon, A. (2004). Grammatical swarm. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 163–174.

O'Neill, M. and Brabazon, A. (2006). Grammatical swarm: The generation of programs by social programming. *Natural Computing*, 5(4):443–462.

Parsopoulos, K., Plagianakos, V. P., and Magoulas, G. D. (2001). Stretching technique for obtaining global minimizers through particle swarm optimization. In *Proceedings of Particle Swarm Optimization Workshop*, pages 22–29.

Penga, Y. and Lua, B.-L. (2013). A hierarchical particle swarm optimizer with latin sampling based memetic algorithm for numerical optimization. *Applied Soft Computing*, 13(5):2823–2836.

Prasad, R. (2004). *Genetic Algorithms: Genesis of Stock Evaluation*. Aston University.

Rapaic, R. M., Kanovic, Z., and Jelicic, Z. D. (2009). A theoretical and empirical analysis of convergence related particle swarm optimization. *WSEAS Transactions on Systems and Control*, 4:541–550.

Sarlin, P. (2014). A weighted som for classifying data with instance-varying importance. *International Journal of Machine Learning and Cybernetics*, 5(1):101–110.

Singh, G., Grandhi, R. V., and Stargel, D. S. (2010). Modified particle swarm optimization for a multimodal mixed-variable laser peening process. *Structural and Multidisciplinary Optimization*, 42(5):769–782.

van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., and Srinivasa, R. (2006). A global sensitivity analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology*, 324:10–23.

Xiaoac, Y., Jin, X., Fengbin, L., and Shouyang, W. (2014). Ensemble anns-pso-ga approach for day-ahead stock e-exchange prices forecasting. *International Journal of Computational Intelligence Systems*, 7(2):272.

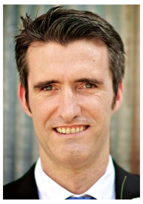
Zambrano-Bigiarini, M. and Rojas, R. (2013). A model-independent particle swarm optimisation software for model calibration. *Environ. Model. Softw.*, 43:5–25.

Zhang, F. and Liao, Z. (2014). Gold price forecasting based on rbf neural network and hybrid fuzzy clustering algorithm. *Lecture Notes in Electrical Engineering*, 241:73–84.

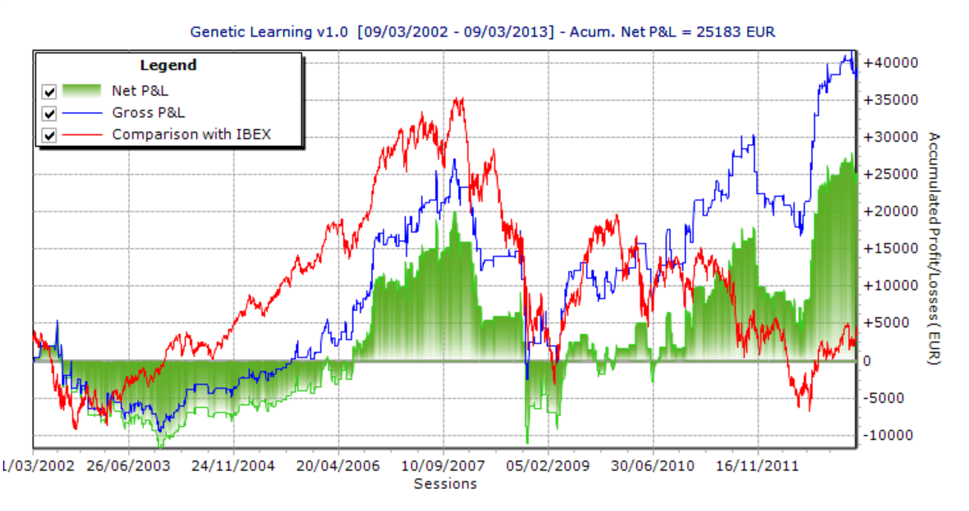
Authors' Information

Nuria Gómez Blas *- Dept. Organización y Estructura de la Información, Escuela Univesitaria de Informática, Universidad Politécnica de Madrid, Crta. de Valencia km. 7, 28031 Madrid, Spain; e-mail: ngomez@eui.upm.es*
Major Fields of Scientific Research: Bio-inspired Algorithms, Natural Computing

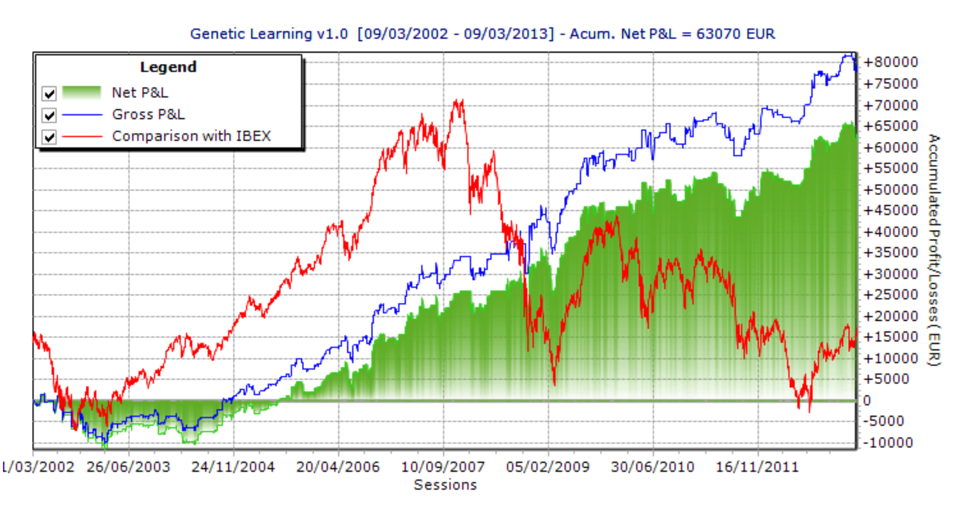
Luis Fernando de Mingo López *- Dept. Organización y Estructura de la Información, Escuela Univesitaria de Informática, Universidad Politécnica de Madrid, Crta. de Valencia km. 7, 28031 Madrid, Spain; e-mail: lfmingo@eui.upm.es*
Major Fields of Scientific Research: Artificial Intelligence, Social Intelligence



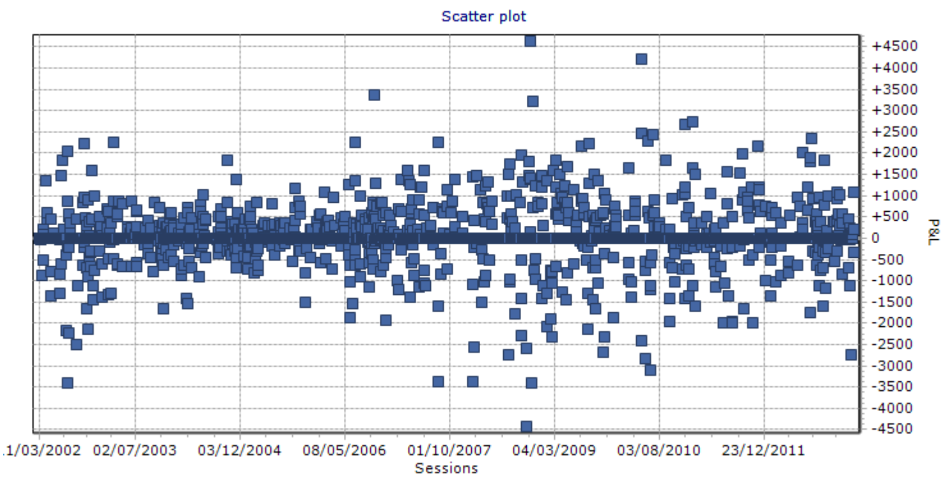
Alberto Arteta *- Dept. Tecnología Fotónica, ETSI Telecomunicación, Universidad Politécnica de Madrid, Avenida Complutense 30, Ciudad Universitaria, 28040 Madrid, Spain; e-mail: m.muriel@upm.es*
Major Fields of Scientific Research: Theoretical Computer Science, Microwave Photonics



a) Accumulated profit/losses, net and gross

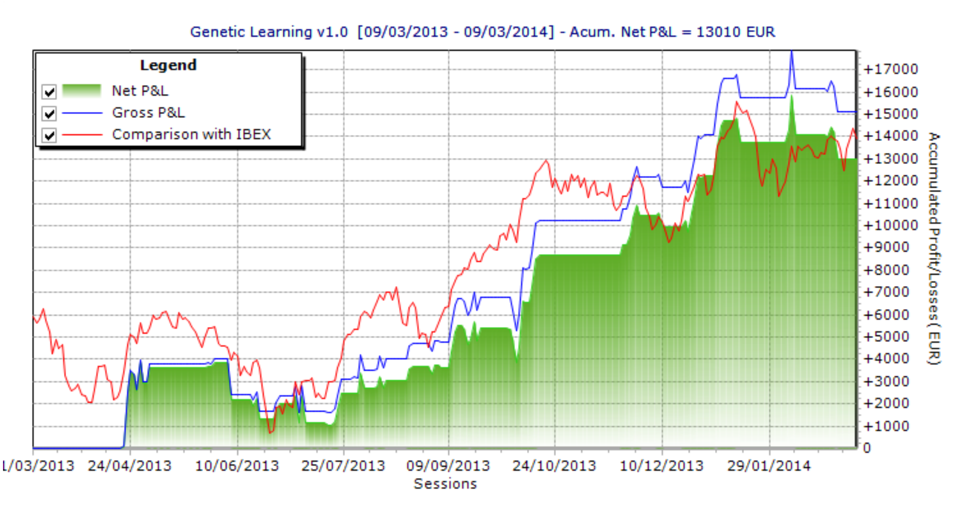


b) Scattered plot of winning/losing sessions

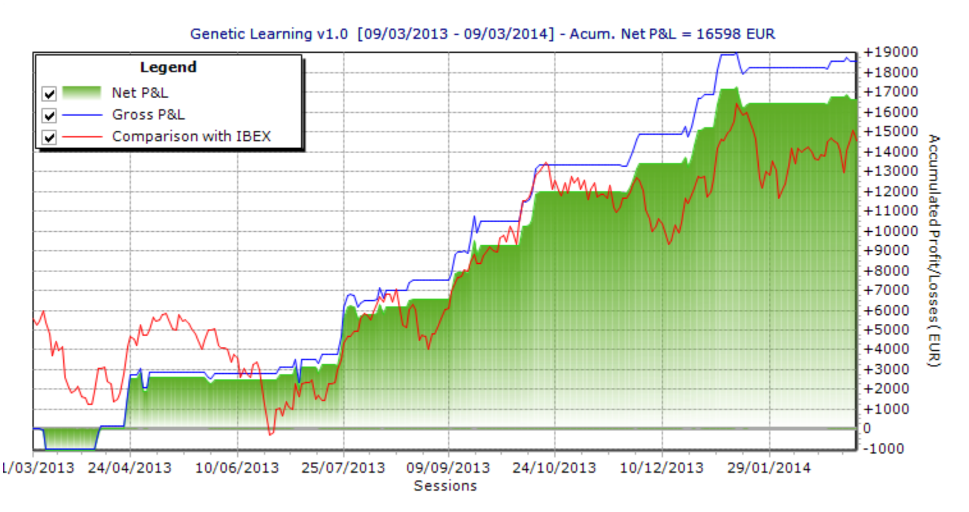


c) Maximum drawdown (worst session losses)

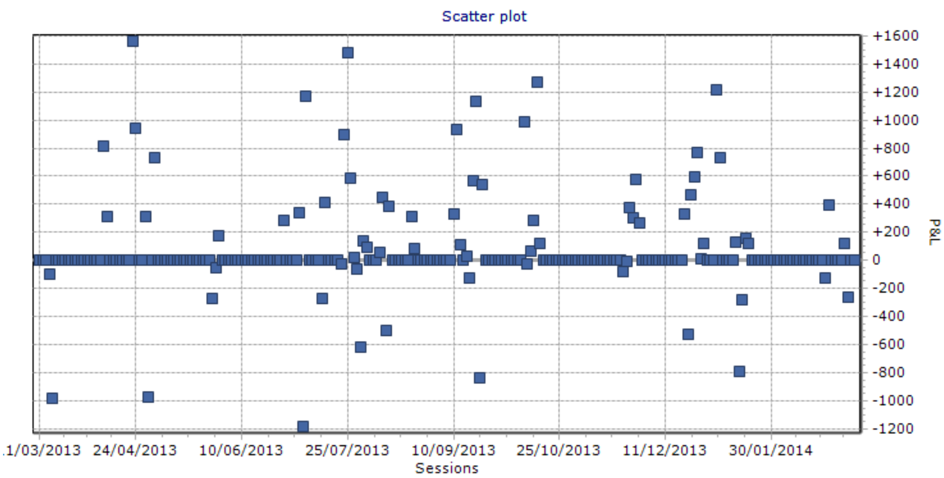
Figure 1: Neural network training performance. The training data set is obtained from `www.tradingmotion.com` and it starts from 2002 till present with a period of 30 minutes.



a) Accumulated profit/losses, net and gross



b) Scattered plot of winning/losing sessions



c) Maximum drawdown (worst session losses)

Figure 2: Neural network testing performance. The system has been cross-validated with real stock market indicators

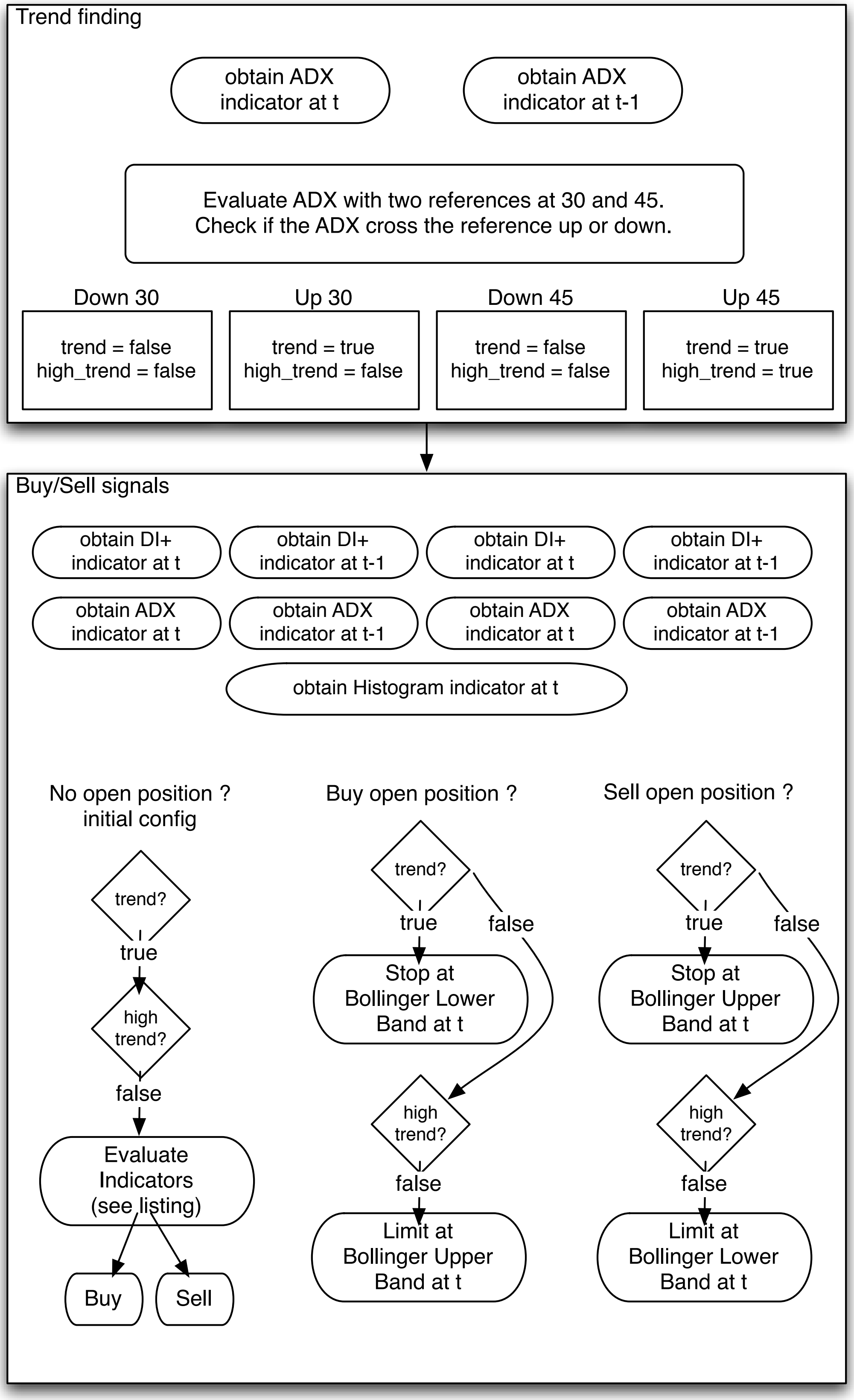
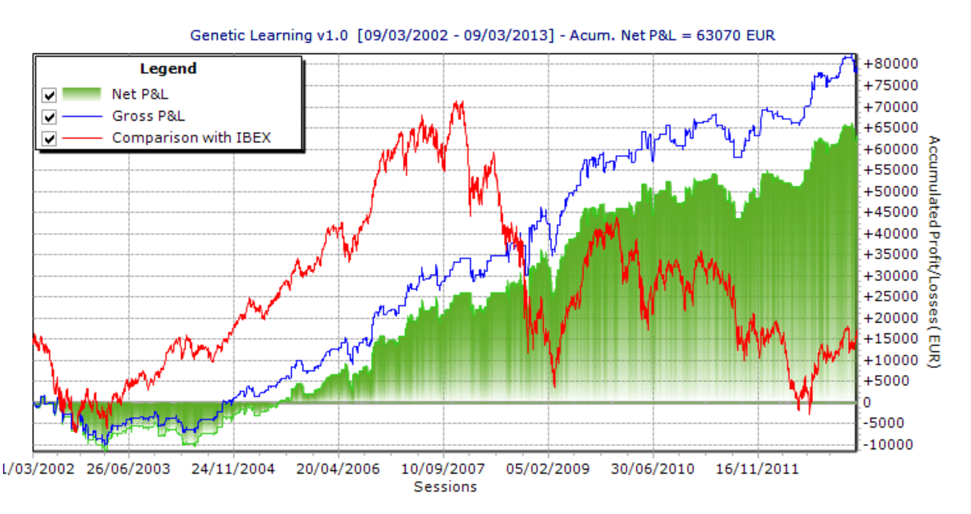
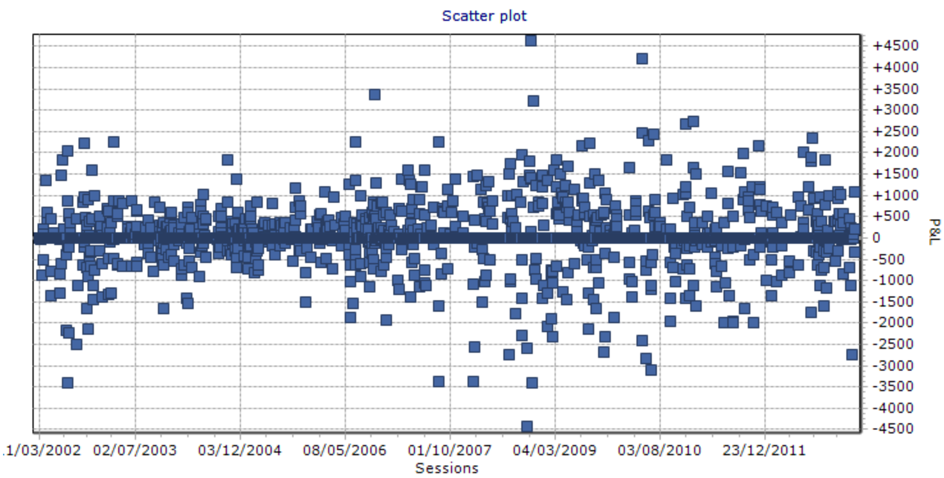


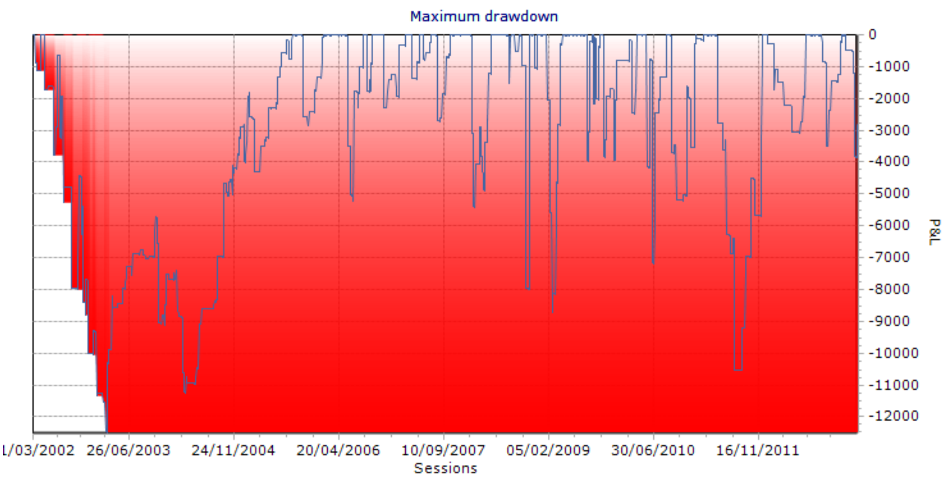
Figure 3: Data flow corresponding to the obtained algorithm using grammatical swarm, see section to find the C++ listing.



a) Accumulated profit/losses, net and gross

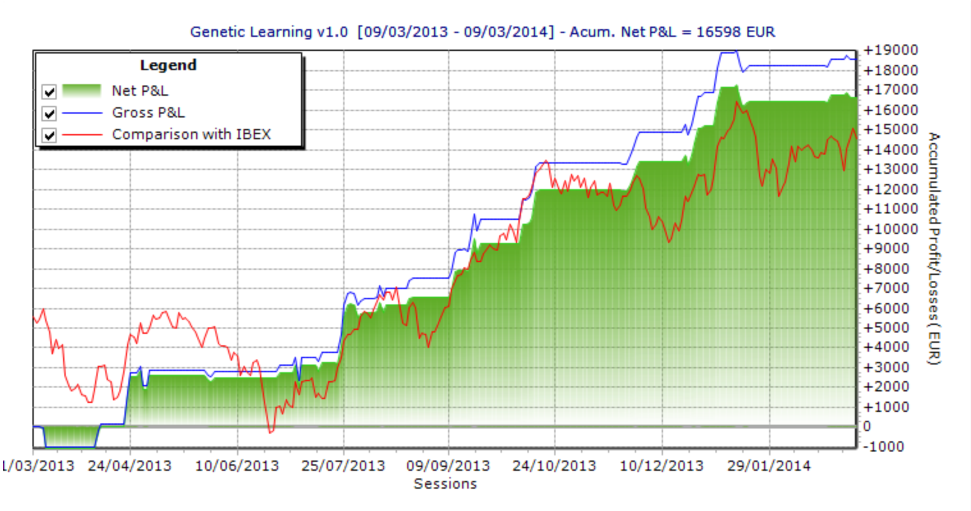


b) Scattered plot of winning/losing sessions

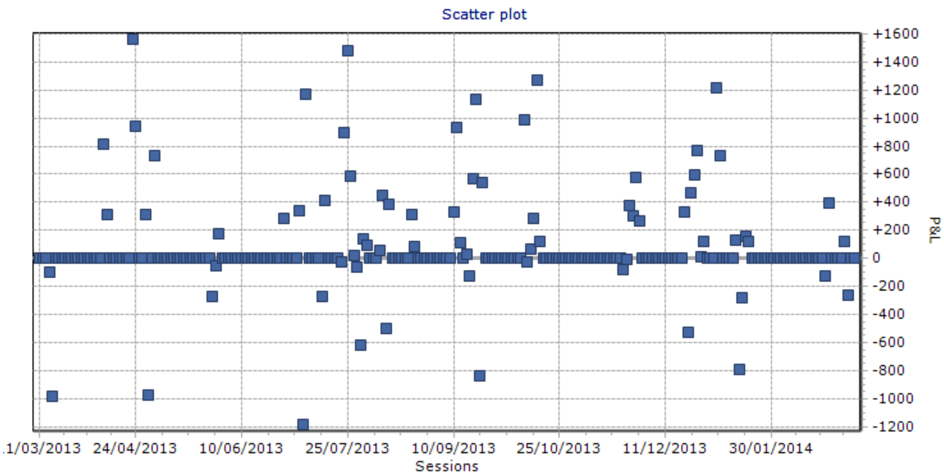


c) Maximum drawdown (worst session losses)

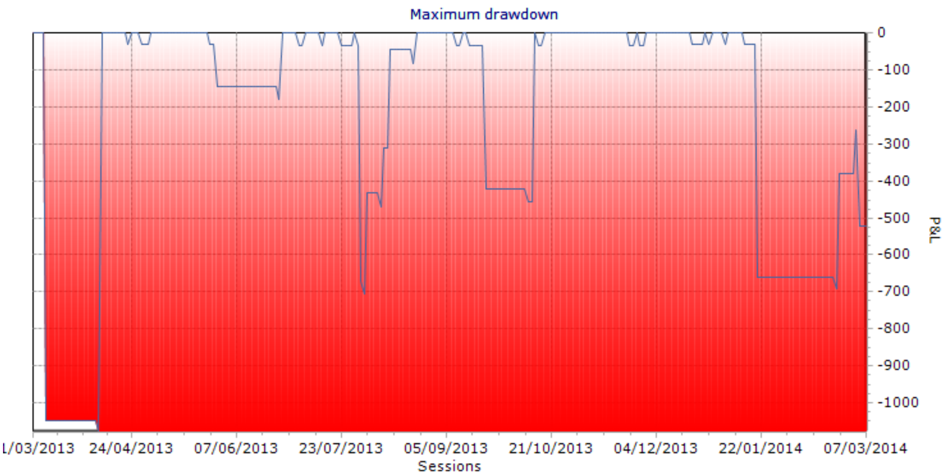
Figure 4: Grammatical swarm training performance. The system has been trained with historical data available at `www.tradingmotion.com` with a period of 30 minutes, corresponding to real stock market indicators



a) Accumulated profit/losses, net and gross



b) Scattered plot of winning/losing sessions



c) Maximum drawdown (worst session losses)

Figure 5: Grammatical swarm testing performance. Data testing is performed with non trained data set corresponding to real values.

Search: ibex Show 25 entries First Previous 1 2 Next Last EUR

★	System	Product		Developer	Type	Start	Total P/L	Annual ROI	Winning Sessions	Profit Factor	Worst Drawdown	Suggested / Required
☆	Trend A1 Ibex 11'	IX		AutoTradingBot	I	1/2001	258091 €	+63.7%	50.3%	1.65	-9442 €	30000 € 2200 €
☆	Innova 221 Ibex	IX		TodoBolsa	I	1/2006	90985 €	+53.4%	56.2%	1.70	-5288 €	20000 € 2200 €
☆	Hercules_1151 Ibex	IX		TodoBolsa	I	1/2001	214284 €	+52.8%	49.9%	1.44	-9794 €	30000 € 3000 €
☆	LAG_IBEX_3	IX		Luis Antón	I	1/2001	301176 €	+49.5%	48.5%	1.37	-14242 €	45000 € 2200 €
☆	Intr Break 10' Ibex p.p. E	IX		TradingMotion	I	1/2001	198760 €	+49.0%	49.7%	1.38	-9989 €	30000 € 2700 €
☆	Intr Break 10' Ibex p.p. A	IX		TradingMotion	I	1/2001	197545 €	+48.7%	49.6%	1.38	-10598 €	30000 € 2700 €
☆	Cibeles 1051 Ibex	IX		SistemasTrading	I	1/2001	208018 €	+43.9%	46.6%	1.44	-11800 €	35000 € 2700 €
☆	MagicBreak Ibex 15' V2	IX		Joel Galí	I	1/2001	156234 €	+38.5%	49.4%	1.62	-9346 €	30000 € 3300 €
☆	Intr Break 10' Ibex p.p. B	IX		TradingMotion	I	1/2001	204243 €	+37.7%	49.8%	1.40	-15123 €	40000 € 2700 €
☆	NinjaCrea 10' Ibex	IX		CreaTrading	S	4/2006	149685 €	+36.2%	56.6%	1.65	-8929 €	50000 € 11200 €
☆	Intr Tauro 9' Ibex	IX		SistemasTrading	I	1/2001	236609 €	+35.0%	49.0%	1.38	-16876 €	50000 € 2200 €
☆	Velazquez Ibex 17'	IX		Javier González-Barros	S	1/2001	234830 €	+34.7%	58.5%	1.43	-13191 €	50000 € 11200 €
☆	Barrix Ibex 17'	IX		Javier González-Barros	S	1/2001	255280 €	+34.3%	59.0%	1.45	-14352 €	55000 € 11200 €
☆	Univer Ibex 15'	IX		Javier González-Barros	S	1/2001	202590 €	+27.2%	47.5%	1.45	-17426 €	55000 € 11200 €

Figure 6: Ranking of best automatic systems forecasting Spanish stock IBEX index. Information can be check it at `www.tradingmotion.com`. Note that are sorted by annual Return On Investment (ROI).

PROCESSING SETS OF CLASSES' LOGICAL REGULARITIES

Anatoliy Gupal, Maxim Novikov, Vladimir Ryazanov

***Abstract**: The paper considers two methods for processing sets of logical regularities of classes (LRC) found by training samples analysis. The first approach is based on the minimization of logical descriptions of classes. As a result of solving the problem of linear discrete optimization, the shortest logic description of each class is found. Each training object satisfies at least to one LRC of found irreducible subset of logical regularities. The second approach is based on the clustering of the set of LRC and selecting standards of derived clusters. The clustering problem is reduced to the clustering of representations of LRC set. Here each LRC is represented in the form of binary vector with different informative weight. A modification of the known method of "variance criterion minimization" for the case where the objects have different information weights is proposed. We present the results of illustrative experiments.*

***Keywords**: classification, logical regularity of class, feature, clustering*

***ACM Classification Keywords**: I.2.4 Artificial Intelligence Knowledge Representation Formalisms and Methods – Predicate logic; I.5.1 Pattern Recognition Models – Deterministic, H.2.8 Database Applications, Data mining*

Introduction

Methods of classification by precedents got now widely fame and spread in various fields of science, production and social life. There are many practical problems where the initial data (training samples) have the form $\{w_i, \mathbf{x}_i, i = 1,2,...,m\}$, $\mathbf{x}_i = (x_{i1}, x_{i2},..., x_{in})$ is a feature description of object, $w_i \in \{1,2,...,l\}$ is a number of "class" of the object. We will further assume that the features of the object are real values, the object itself and its feature descriptions will be identified, training sample comprises representatives of each class, the training sample is noncontradictory (there are no equal objects in different classes). The main problem is to construct algorithm A that for any new feature description $\mathbf{x}$ will calculate its class: $w = A(\mathbf{x})$. Currently, there are a lot of different approaches and specific algorithms for solving classification problems based on different principles and ideas. Statistical algorithms are based on probabilistic hypotheses about the structure of classes and suggest the existence of a priori probabilities of classes and conditional probability distributions for each class of objects [Duda et al, 2000]. Methods based on the construction of separating surfaces, find the optimal separating surfaces for training data (linear, polynomial, piecewise-linear, linear combinations of potential functions, etc.) and use them for further classification [Duda et al, 2000; Vapnik, 1995]. Neural network approaches are the methods to find the optimal superposition of threshold functions and are attempt to simulate the human thinking [Wasserman, 1992]. Logical approaches (models of partial precedent [Zhuravlev, 1978; Zhuravlev et al, 2006] and decision trees [Quinlan, 1993]) are natural generalization of the binary representations of functions in disjunctive normal form. Many approaches are in fact hybrid and combine several different principles. The present work is related to models of a partially precedent approach. Along with the creation of the classifier A on the training set, practical user is interested in illustrative communication by analytical formulas between baseline characteristics and classes, in descriptions of the classes themselves. In this regard, the logical approaches have certain advantages. They can be used to construct as recognition algorithms and also analytical description of classes. Article is devoted to the second part. The LRCs may be similar and even equal. LRC may be some degenerate solution, the number of calculated LRC can be quite large. Practical user is interested in the same small number of LRC, but different and informative. This situation undoubtedly requires the development and use of various means of processing LRC.

Recognition algorithms based on the weighted voting over system of logical regularities

Let us consider the problem of recognition by precedents (supervised classification) in the following standard statement [Zhuravlev, 1978]. We believe that given a set M of objects $\mathbf{x}$. Each object is given in terms of feature values $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}, x_i \in R$, and $M = \{\mathbf{x}\} = \bigcup_{i=1}^{l} K_i, K_i \cap K_j = \varnothing, i, j = 1,2,\ldots,l, i \neq j$. Subsets K_i are called classes. The initial information about M and its partition to classes given in the form of training sample $X = \{\mathbf{x}\} \subseteq M$, containing representatives of all classes. Training sample is consistent, i.e. not containing equal objects of different classes. Required to set up an algorithm A that classifies any object $\mathbf{x}$ to one of the classes, or refuse to classify the object.

We will write and consider $(a \le x)$ the expression equal to one when it is running, and zero otherwise. Next, we use the following definition LRC [Ryazanov, 2007].

Definition 1. The predicate

$$P^{\mathbf{a},\mathbf{b},\Omega_1,\Omega_2}(\mathbf{x}) = \mathop{\&}_{i \in \Omega_1} (a_i \le x_i) \mathop{\&}_{i \in \Omega_2} (x_i \le b_i)$$

is a logical regularity of class (LRC) $K_t, t = 1,2,\ldots,l$, if the following conditions are satisfied

1. $\exists \mathbf{x}_j \in X \cap K_t : P^{\mathbf{a},\mathbf{b},\Omega_1,\Omega_2}(\mathbf{x}_j) = 1$.

2. $\forall \mathbf{x}_j \notin X \cap K_t : P^{\mathbf{a},\mathbf{b},\Omega_1,\Omega_2}(\mathbf{x}_j) = 0.$

3. $F\left(P^{\mathbf{a},\mathbf{b},\Omega_1,\Omega_2}(\mathbf{x})\right) = \text{extr}\, F\left(P^{\mathbf{a}^*,\mathbf{b}^*,\Omega_1^*,\Omega_2^*}(\mathbf{x})\right)$, where $\mathbf{a}^*, \mathbf{b}^* \in R^n, \mathbf{a}^* \le \mathbf{b}^*, \Omega_1^*, \Omega_2^* \subseteq \{1,2,\ldots,n\}$, F - some criterion of predicate quality.

In the last condition we consider the problem of local maximizing of the criterion $F\left(P^{\mathbf{a},\mathbf{b},\Omega_1,\Omega_2}(\mathbf{x})\right) = \left|\left\{\mathbf{x}_j \in X \cap K_t : P^{\mathbf{a},\mathbf{b},\Omega_1,\Omega_2}(\mathbf{x}_j) = 1\right\}\right|$. Two LRC of class K_t will be called equivalent if their values on the objects of training sample are the same. The set $N_\lambda = \left\{\mathbf{x} : P_\lambda^{\mathbf{a},\mathbf{b},\Omega_1,\Omega_2}(\mathbf{x}) = 1\right\}$ is called as interval for LRC $P_\lambda^{\mathbf{a},\mathbf{b},\Omega_1,\Omega_2}(\mathbf{x})$. Since (for simplicity), we believe that all the features are real, there are a continuum of LRC equivalent for any LRC. For any set of equivalent LRC there is the only minimal LRC. In [Kovshov, 2008] are shown the various approximate and exact methods of finding minimal LRC. One can show that they have the form $P^{\mathbf{a},\mathbf{b}}(\mathbf{x}) = \underset{i \in 1,2,\ldots,n}{\&} (a_i \le x_i \le b_i)$. Thus, we assume that using the training set X for each class K_t is calculated a lot of some LRC: $\mathbf{P}_t = \left\{P_i^{\mathbf{a}_i,\mathbf{b}_i}(\mathbf{x})\right\}$. We put in correspondence the set of intervals $\mathbf{N}_t$ for each set $\mathbf{P}_t$. In [Ryazanov, 2007] presented a method for calculating the coefficients γ_λ in the expression for estimates of the classes in the form $\Gamma_t(\mathbf{x}) = \sum_{P_\lambda \in \mathbf{P}_t} \gamma_\lambda P_\lambda^{\mathbf{a}_\lambda,\mathbf{b}_\lambda}(\mathbf{x})$. Further, two approaches are proposed to processing sets $\mathbf{P}_t$.

Building shortest logical descriptions of classes

In [Zhuravlev et al, 2006] proposed the based on the minimization of logical descriptions of classes method for processing a set of LRC. This approach is similar to the minimization of partially defined Boolean functions. Suppose that for a certain class K_t some system $\mathbf{P}_t$ of LZC was found. From their definition it follows that as an approximation of the characteristic function of the class can take the disjunction of all LRC. This function is equal to 1 on all objects of training of K_t and 0 on all objects of training from other classes.

Definition 2 [Zhuravlev et al, 2006]. Logical description of a class K_t is a function $D_t(\mathbf{x}) = \underset{P_i \in \mathbf{P}_t}{\vee} P_i^{\mathbf{a}_i,\mathbf{b}_i}(\mathbf{x})$.

LRC can be calculated using the recognition system [Zhuravlev et al., 2006]. By analogy with the minimization of partially defined Boolean functions definitions of the shortest and the minimal logical descriptions of classes are introduced.

Definition 3 [Zhuravlev et al, 2006]. Shortest logical description of the class K_t is a function $D_t^s(\mathbf{x}) = \underset{P_i \in \tilde{\mathbf{P}}_t \subseteq \mathbf{P}_t}{\vee} P_i^{\mathbf{a}_i,\mathbf{b}_i}(\mathbf{x})$ where

1. $D_t^s(\mathbf{x}_i) = D_t(\mathbf{x}_i), i = 1, 2, ..., m$
2. $\left|\tilde{\mathbf{P}}_t\right| \to \min\limits_{\tilde{\mathbf{P}}_t \subseteq \mathbf{P}_t}$.

Figures 1 and 2 are examples of the covering of sets of training objects from the class K_t by intervals N_i corresponding to all LRC of $\mathbf{P}_t$ and LRC of $\tilde{\mathbf{P}}_t$. In the example, two marked by circles with a black and a white center classes of two-dimensional objects have been considered.

Here, as an approximation of the characteristic function is sufficient to use 3 logical regularities. The shortest logical description of the class is a result of the solution of integer linear programming problem (here N is the number of all LRC):

$$\sum_{i=1}^{N} y_i \to \min,$$

$$\sum_{i=1}^{N} P_i(\mathbf{x}_j) y_i \geq 1, \forall \mathbf{x}_j \in K_t, y_i \in \{0,1\}.$$

The set of unit values of y_i defines a subset of predicates $\tilde{\mathbf{P}}_t$.

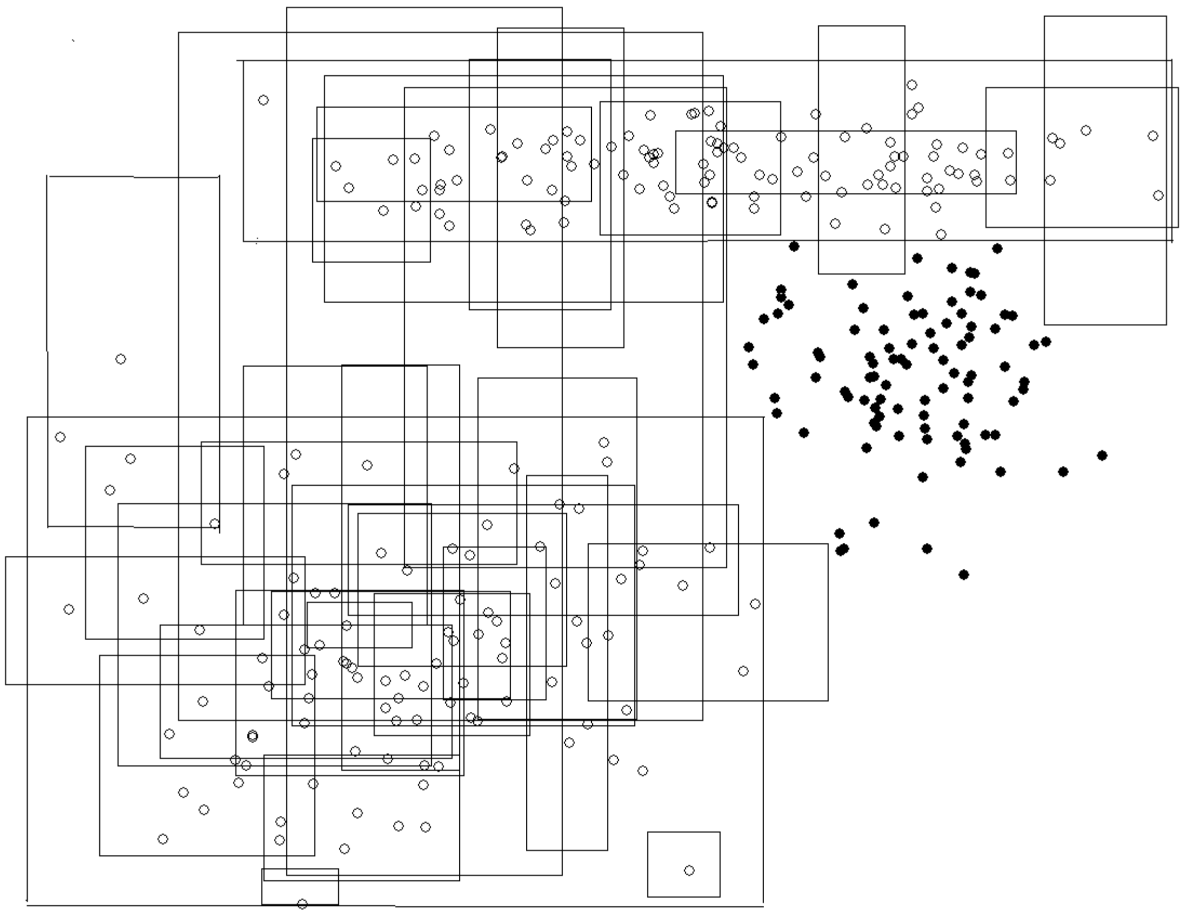


Figure 1. Covering points of the first class of training sample by found intervals from $\mathbf{N}_t$

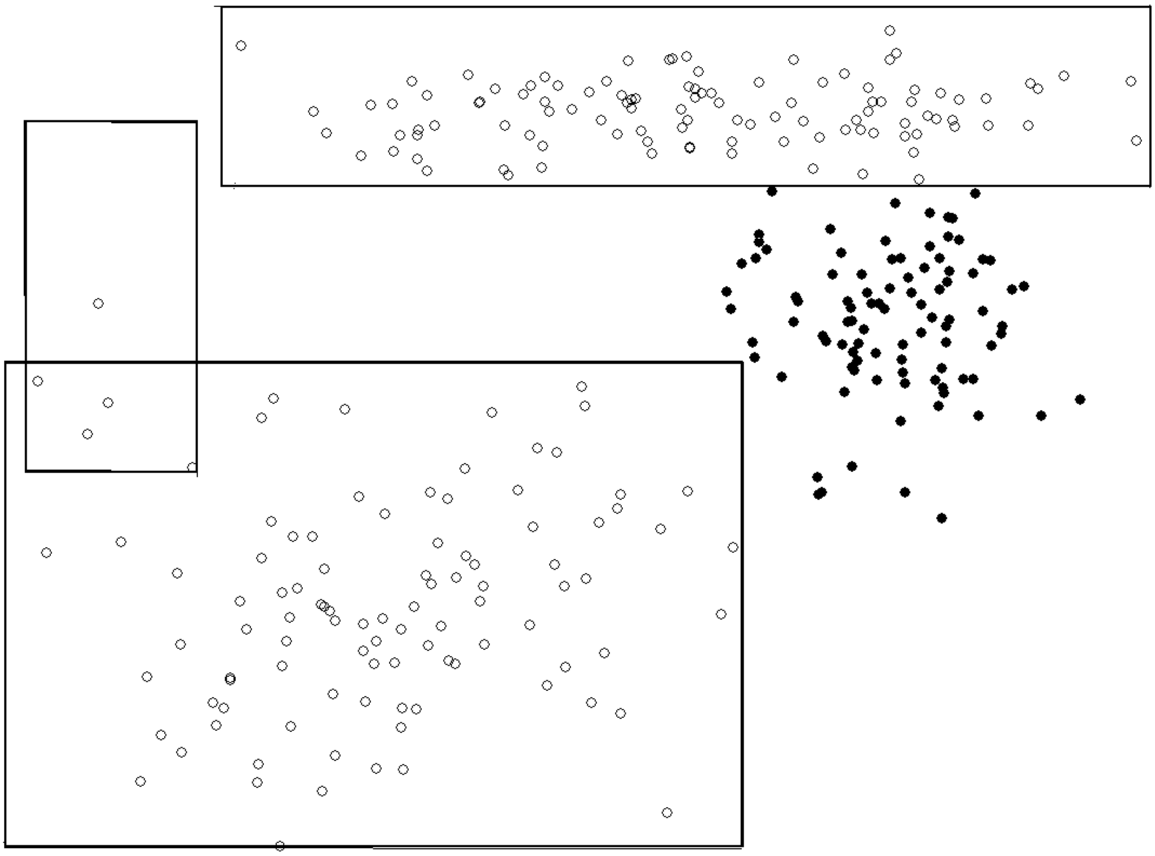


Figure 2. Covering points of the first class of training sample by found intervals from $\tilde{\mathbf{N}}_t$

Processing of sets of LRC using cluster analysis

The second approach to processing sets of LRC is based on cluster analysis. The general idea is as follows. When processing a set of vectors we can solve the problem of clustering for 2, 3, ... clusters depending on the prior knowledge or desire. Its "standard" is calculated (for example, the sample mean vector) for each cluster. The resulting system "of standards" is taken as the result of processing of the original set of vectors of precedents. In our case, the objects are clustering functions (LRC). As a result of clustering of the set LRC for l classes and computation for each cluster its standards we get new predicates, which in general can not be in the original set. These predicates are generally "partial" LRC, i.e. they can take the value 1 on a few of objects of other classes. These predicates are "sufficiently" different, they are measured. Thus, by the initial set of LRC we can calculate and evaluate a given number of "sufficiently" different partial logical regularities. Figures 3 and 4 are examples for illustration. On Figure 3 points of a class (the class "circles") are covered by the system of LRC. Figure 4 shows the same example, but is shown only two intervals. Intervals are significantly different and mainly cover a significant number of points of the class under consideration.

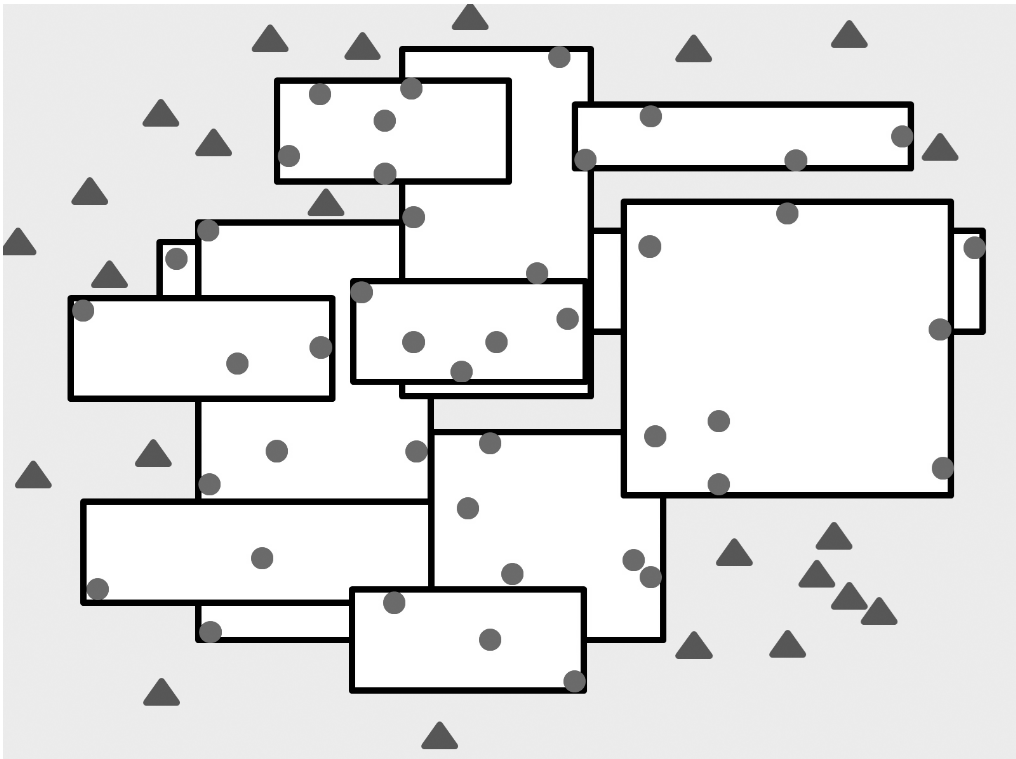


Figure 3. The points covered by a large number of class intervals

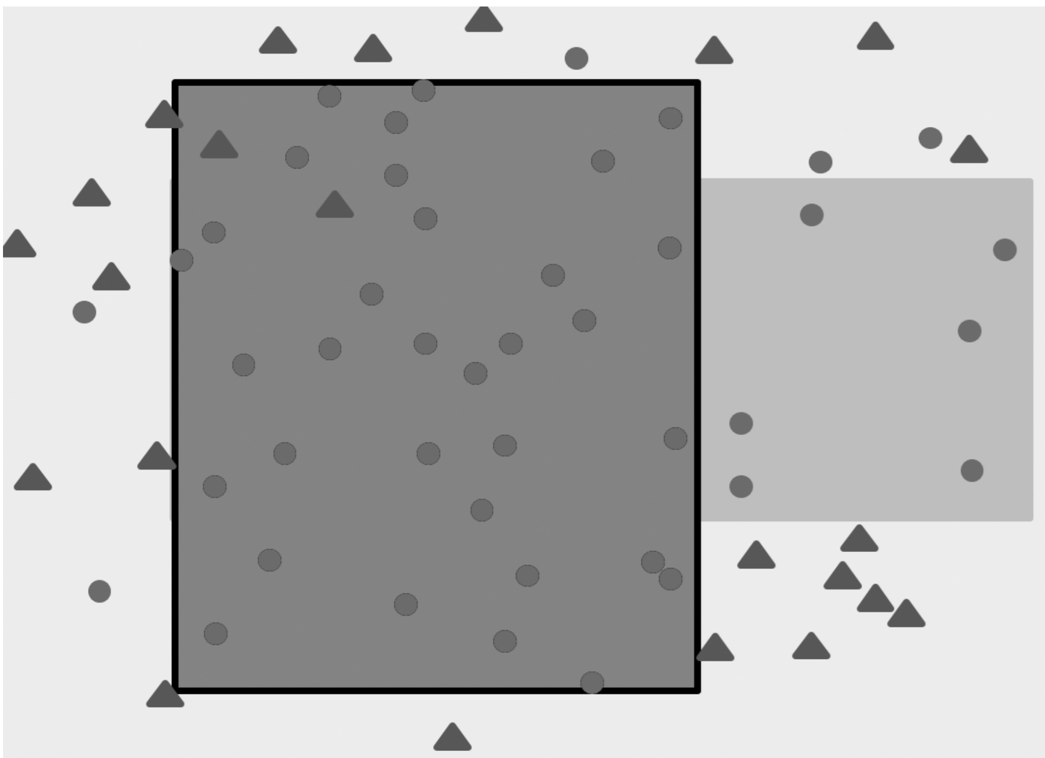


Figure 4. "Substantial" number of class points is covered by two intervals containing perhaps a small number of elements of other classes

To implement this idea we must to create a method of clustering of a set of functions and to calculate the "standard" for a variety of functions that form a cluster. The functions must be weighted. There are two possible ways. The first way is connected with the generalization of existing approaches to clustering the set of vectors to the set of functions. The second way is connected with a one-to-one representation LRC in the form of vectors. In this case, the application of existing clustering methods is possible. The present work is devoted to the second approach.

Put in one-to-one correspondence of each $P_i^{\mathbf{a}_i,\mathbf{b}_i}(\mathbf{x})$ of $\mathbf{P}_t$ the binary vectors as follows: $P_i^{\mathbf{a}_i,\mathbf{b}_i}(\mathbf{x}) \Leftrightarrow \mathbf{z}_i = (\mathbf{z}_{i1}, \mathbf{z}_{i2}, ..., \mathbf{z}_{ih}), \mathbf{z}_{ij} \in \{0,1\}, j = 1,2,...,h$. Here $h = |K_t|$, the vector $\mathbf{z}_i$ marks the original training objects from the class K_t in which the predicate $P_i^{\mathbf{a}_i,\mathbf{b}_i}(\mathbf{x})$ is unity. The weight of each vector $\mathbf{z}_i$ (and corresponding LRC $P_i^{\mathbf{a}_i,\mathbf{b}_i}(\mathbf{x})$) equals to a fraction of class objects K_t for which LRC is equal to 1. Thus, the original problem is reduced to the clustering of the set of binary vectors $\mathbf{z}_i = (\mathbf{z}_{i1}, \mathbf{z}_{i2}, ..., \mathbf{z}_{ih})$ with known weights $y(\mathbf{z}_i)$ and calculation of the mean of each cluster. Further, each of the sample mean is associated with some in general partial LRC. As a basic clustering method we take the method based on the minimization of variance criterion. Suppose we have fixed the number of clusters l.

Clustering on l clusters by minimizing the variance criterion is formulated in the following way:

$$J(\mathrm{K}) = \sum_{i=1}^{l} \sum_{\mathbf{z}_t \in K_i} y_t \left\| \mathbf{z}_t - \mathbf{m}_i \right\|^2 \to \min_{\mathrm{K}} \tag{1}$$

where $y_t = y(\mathbf{z}_t)$, $\mathrm{K} = \bigcup_{i=1}^{l} K_i, K_i \cap K_j = \varnothing, i \neq j, i, j = 1,2,...,l$, $\mathbf{m}_i = \dfrac{\sum_{\mathbf{z}_t \in K_i} y_t \mathbf{z}_t}{\sum_{\mathbf{z}_t \in K_i} y_t}$.

In a standard clustering algorithm we have $y_i = 1, i = 1,2,...,m$. It is known that in this case, while minimizing the criterion $J(\mathrm{K}) = \sum_{j=1}^{l} \sum_{\mathbf{x}_i \in K_j} \left\| \mathbf{z}_i - \mathbf{m}_j \right\|^2 = \sum_{j=1}^{l} J_j(K_j)$, the local optimality condition for $\mathrm{K} = \{K_1, K_2, ..., K_l\}$ is the implementation of inequalities

$$\frac{n_i}{(n_i - 1)} \left\| \hat{\mathbf{z}} - \mathbf{m}_i \right\|^2 - \frac{n_j}{(n_j + 1)} \left\| \hat{\mathbf{z}} - \mathbf{m}_j \right\|^2 \leq 0 \tag{2}$$

for any pair K_i, K_j, and any $\hat{\mathbf{z}} \in K_i$. This means that any movement of the object of the cluster to which it belongs to any other cluster does not lead to a reduction of the dispersion criterion. We can prove that in the general case there is an analogue (3) of condition (2) for the task (1), when there are not unitary weights. For simplicity, we write $\sum y$ instead of $\sum_{\mathbf{z}_t \in K_i} y_t$, and $\sum y\mathbf{z}$ instead of $\sum_{\mathbf{z}_t \in K_i} y_t \mathbf{z}_t$.

We have the expressions $\mathbf{m}_i^* = \mathbf{m}_i - \frac{\hat{y}(\hat{\mathbf{z}} - \mathbf{m}_i)}{(\sum y - \hat{y})}$ and $\mathbf{m}_j^* = \mathbf{m}_j + \frac{\hat{y}(\hat{\mathbf{z}} - \mathbf{m}_j)}{(\sum y + \hat{y})}$ when transferring $\hat{\mathbf{z}}$ from K_i to K_j .

Then $K_i \to K_i^* = K_i \setminus \{\hat{\mathbf{z}}\}, K_j \to K_j^* = K_j \cup \{\hat{\mathbf{z}}\}$ and

$$\frac{\sum_{K_i} y\hat{y}}{(\sum_{K_i} y - \hat{y})} \left\| \hat{\mathbf{z}} - \mathbf{m}_i \right\|^2 - \frac{\sum_{K_j} y\hat{y}}{(\sum_{K_j} y + \hat{y})} \left\| \hat{\mathbf{z}} - \mathbf{m}_j \right\|^2 \le 0 \tag{3}$$

As a result, the vector $\mathbf{m}_i = (m_{i1}, m_{i2}, ... m_{ih})$ is calculated for each cluster. By construction, $m_{ij} \in \{0, \alpha_{i1}, \alpha_{i2}, ..., \alpha_{iu}\}, 0 \le \alpha_{i\sigma} < \alpha_{i,\sigma+1} \le 1$ and values $\alpha_{i\sigma}$ are calculated by the obtained clustering. The standard of each cluster will be assumed the Boolean vector $\mathbf{b}_i = (b_{i1}, b_{i2}, ..., b_{ih}), b_{ij} \in \{0,1\}$, where $b_{ij} = \begin{cases} 1, & m_{ij} \ge \theta_i, \\ 0, & \text{otherwise.} \end{cases}$ Here, θ_i is chosen from a finite set $D_i = \{0, \alpha_{i1}, \alpha_{i2}, ..., \alpha_{iu}\}$. The vector $\mathbf{b}_i$ is in correspondence to each selection of θ_i , the selection of the optimal value of the vector is carried out by solving the problem of one-dimensional optimization $\Phi(P(\mathbf{x}, \theta_i)) \to \max_{\theta_i \in D_i}$. Note that, in general, partial logical regularities of classes are correspondent to the $\mathbf{b}_i$.

There are different criteria Φ of quality of partial logical regularities of class, which correspond to a choice of θ_i . An example of a criterion $\Phi(P(\mathbf{x}, \theta))$ may be the criterion $\Phi(P(\mathbf{x}, \theta)) = \sqrt{p_i} - \sqrt{n_i}$ [Cohen et al, 1999], where p_i is a number of training objects of the class K_i in which the corresponding to the selected θ_i predicate $P(\mathbf{x})$ is performed, n_i is the number of training objects of other classes, for which the predicate is executed. There are various other evaluation criteria for $P(\mathbf{x})$, that take into account the number of sampling objects in the class and the number of other training objects.

The experimental results on model and practical problems

Practical experiment was performed on problem „breast” [Mangasarian et al, 1990]. This problem of recognition of breast cancer by precedents contained 218 first-class objects (benign tumor) and 126 objects of the second class (malignant neoplasm). Each object is described by 9 discrete features which are denoted as $X1, X2, ..., X9$. For the second class with an approximate algorithm [Zhuravlev et al, 2006] was found 17 logical regularities and the shortest logical description of the class from 7 conjunctions. Below in Table 1, the shortest logical description is represented on the left side. For each conjunction, its serial number from the overall list LRC and share objects of the second training class that have been met are given.

Table 1. Comparison of clustering standards on data „breast”

	$\Phi(P(\mathbf{x},\theta))/K_2$	n_i	Logical regularities
(3.5<=X5) (4.5<=X6) (X8<=6.5) {60, 0.28} V (6.5<=X4) {61, 0.40} V (6.5<=X3) {77, 0.48}	0.84	8	$(4 \le x_2 \le 10)(3 \le x_3 \le 10)(2 \le x_5 \le 10)(2 \le x_7 \le 10)$
V (X6<=4.1)(4.0<=X8) {78, 0.14} V (6.5<=X1)(3.3<=X6) {79, 0.54}	0.78	5	$(3 \le x_1 \le 10)(2 \le x_3 \le 10)(2 \le x_5 \le 10)(4 \le x_6 \le 10)$
V (2.0<=X3)(4.5<=X7) {81, 0.67} V (2.0<=X1)(5.8<=X9) {83, 0.10}	0.67	4	$(2 \le x_3 \le 10)(2 \le x_5 \le 10)(5 \le x_7 \le 10)(1 \le x_9 \le 8)$

In the right part of Table 1, there are shown the standards of resulting clustering of 17 logical regularities of second class on 3 clusters by the use of criterion [Cohen et al, 1999] for construction of standards. There are shown the values of quality and power of obtained clusters. The problem of minimizing the number of features in the standards was not considered. The arrows indicate the LRC of the shortest logical description that are closest to the calculated standards (here the Hamming distance between the 126-dimensional binary vectors was used).

Conclusion

In this paper we have proposed two approaches to processing a set of logical regularities of class based on finding of the shortest logical descriptions and methods of cluster analysis. An illustrative comparison of these approaches on the same medical problem was conducted. Both approaches are useful tools for analyzing sets of LRCs. The first approach gives us a method to select practically a small number of LRC of their complete list, using the principle of minimum complexity class descriptions. In the second approach, new 1, 2, 3, ... standard LRCs (generally partial LRCs) representing the original set are calculated based on the existing set of LRCs. The number of standards is equal to the number of clusters, which is set by the user. Of interest is the creation and use of other models of clustering LRC sets, as well as the use of different criteria and their quality evaluation.

Acknowledgements

This work was supported by the Program of the Presidium of RAS №15 „Information, control and intelligent technologies and systems", Program №2 of Mathematical Sciences Department of RAS, RFBR № 14-01-90413 Ukr_a, 14-01-90019 Bel_a, 14-01- 00824 a, 13-01-12033 ofi m, 13-01-90616 Arm_a, 12-01-00912-a, 15-01-05776-a.

Bibliography

[Cohen et al, 1999] William W. Cohen and Yoram Singer Simple, Fast, and Effective Rule Learner, AAAI/IAAI 1999: 335-342.

[Duda et al, 2000] Duda, R. O., Hart, P. E., and Stork, D. G. Pattern Classification. John Wiley and Sons, 2nd edition, 2000

[Kovshov et al., 2008] Kovshov N.V., Moiseev V.L., and Ryazanov V.V. “Algorithms for finding logical regularities in pattern recognition problems”, Computational Mathematics and Mathematical Physics, M .: Nauka. T.48, 2008, N 2, pp. 329-344.

[Mangasarian et al, 1990] Mangasarian O. L., Wolberg W.H.: "Cancer diagnosis via linear programming", SIAM News, Volume 23, Number 5, September 1990, pp 1 - 18.

[Quinlan, 1993] Quinlan, J. R. C4.5: Programs for Machine Learning. Morgan Kaufmann, 1993

[Ryazanov, 2007] Ryazanov V.V. “Logical regularities in pattern recognition (parametric approach)”, Computational Mathematics and Mathematical Physics, T.47, №10, 2007, p.1793-1808

[Vapnik, 1995] Vapnik, V.:The Nature of Statistical Learning Theory. Springer-Verlag, 1995.

[Wasserman, 1992] Wasserman, F. Neural Computing. Theory and Practice. - M.: Mir, 1992. - 240 p.

[Zhuravlev et al, 2006] Yu.I.Zhuravlev, V.V.Ryazanov, O.V.Senko, Recognition. Mathematical methods. Software system. Practical applications. Izd.vo "Fazis", Moscow, 2006, 178 pp.

[Zhuravlev, 1978] Yu.I.Zhuravlev, On the algebraic approach to solving the problems of recognition and classification. Problems of Cybernetics. M .: Nauka, 1978. Issure.33, pp. 5-68.

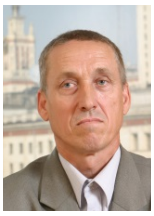
Authors' Information

Anatoliy Gupal *– Corresponding Member of the NAS of Ukraine, Heard of Department; V. M. Glushkov Institute of Cybernetics, Ukraine, 40 Acad. Glushkova Ave., 03680, Kyiv; e-mail: gupal_anatol@mail.ru*

Major Fields of Scientific Research: Bioinformatics, Bayesian network, Markov model

Maxim Novikov *– Junior Engineer at Samsung RnD Institute Russia, Russia, 127018 Moscow, Dvintsev street,12/1; e-mail: maxim.s.novikov@gmail.com*

Major Fields of Scientific Research: Pattern recognition, Data mining, Signal processing



Vladimir Ryazanov *– Heard of Department; Institution of Russian Academy of Sciences Dorodnicyn Computing Centre of RAS, Russia, 119991 Moscow, Vavilov's street, 40; e-mail:* rvv@ccas.ru

Major Fields of Scientific Research: Pattern recognition, Data mining, Artificial Intelligence

CONSTRUCTING AN OPTIMAL INVESTMENT PORTFOLIO BY USING FUZZY SETS THEORY

Yuri Zaychenko, Inna Sydoruk

Abstract: The problem of constructing an optimal portfolio of securities under uncertainty was considered.

The global market crisis of recent years has shown that the existing theory of optimization of investment portfolios and forecasting stock indices exhausted and the revision of the basic theory of portfolio management is strongly needed. Therefore the fuzzy sets theory was used for getting an optimal portfolio.

In this paper direct, dual and multicriteria problems with the use of triangular membership functions work were considered. The problem of portfolio optimization during the time period also was described in this article. In direct task we define structure of a portfolio, which will provide the maximum profitableness at the set risk level. In dual task we define structure of a portfolio, which will provide the minimum risk level at the set level of critical profitableness. In multicriteria problem we simultaneously maximize profitability and minimize risk level. The input data for the optimization system were predicted by using the Fuzzy Group Method of Data Handling (FGMDH). The optimal portfolios for assets were determined. The comparative analysis of optimal portfolios obtained by different methods and approaches was fulfilled.

***Keywords**: membership function, fuzzy sets theory, optimal portfolio, investments, stock securities, fuzzy number, FGMDH*

***ACM Classification Keywords**: G.1.0 Mathematics of Computing– General – Error analysis; G.1.6 Mathematics of Computing – Numerical Analysis – Optimization - Gradient methods, Least squares methods; I.2.3 Computing Methodologies - Artificial Intelligence - Uncertainty, "fuzzy", and probabilistic reasoning.*

Introduction

The problem of investment in securities arouse with appearance of the first stock markets. The main objective of portfolio investment is to improve the investment environment, giving securities such investment characteristics that are only possible in their combination. Careful processing and accounting of investment risks have become an integral and important part of the success of each company. However, more and more companies have to make decisions under uncertainty, which may lead to unintended consequences and, therefore, undesirable results. Particularly serious consequences may have the wrong decisions at long-term investments. Therefore, early detection, adequate and the most accurate assessment of risk is one of the biggest problems of modern investment analysis.

Historically, the first and the most common way to take account of uncertainty is the use of probability theory. The beginning of modern investment theory was in the article H. Markowitz, "Portfolio Selection", which was released in 1952. In this article mathematical model of optimal portfolio of securities was first proposed. Methods of constructing such portfolios under certain conditions are based on theoretical and probabilistic formalization of the concept profitability and risk. For many years the classical theory of Markowitz was the main theoretical tool for optimal investment portfolio construction, after which most of the novel theories were only modifications of the basic theory [Burenin, 1998].

However, the global market crisis of recent years has shown that the existing theory of investment portfolio optimization and forecasting stock indices exhausted itself and a revision of the basic theory of portfolio management is strongly needed.

New approach in the problem of investment portfolio construction under uncertainty is connected with fuzzy sets theory. Fuzzy sets theory was created about half a century ago in the fundamental work of Lotfi Zadeh [Zadeh, 1999]. This theory came into use in the economy in the late 70's. By using fuzzy numbers in the forecast parameters decision - making person was not required to form probability estimates.

The application of fuzzy sets technique enabled to create a novel theory of fuzzy portfolio optimization under uncertainty and risk deprived of drawbacks of classical portfolio theory by Markovitz. In this work we use fuzzy sets theory for getting an optimal investment portfolio. Firstly, portfolio optimization problem in this formulation was considered by O.A. Nedosekin [Nedosekin, 2002]. But in his work only direct problem was considered. The investigations were continued by Esfandiyarfard Maliheh. In [Zaychenko & Maliheh, 2008] the direct optimization problem using different membership functions was considered. However, in these studies, the investor can't determine an optimal portfolio during the time period. Therefore, in this study multiobjective optimization problem in which an investor can prefer risk or profitability using weights coefficients and solve portfolio optimization problem at the chosen time period is considered and analyzed.

Problem statement of portfolio optimization

The purpose of the analysis and optimization of an investment portfolio is research in area of portfolio optimization, and also the comparative analysis of structure of the effective portfolios received using the model Markovitz and fuzzy-set model of a share portfolio optimization.

Let us consider a share portfolio from N components and its expected behavior at time interval $[0, T]$. Each of a portfolio component $i = \overline{1, N}$ at the moment T is characterized by its financial profitableness r_i (evaluated at a point T as a relative increase in the price of the asset for the period) [Zaychenko, 2008].

The holder of a share portfolio – the private investor, the investment company, mutual fund – operates the investments, being guided by certain reasons. On the one hand, the investor tries to maximize the profitableness. On the other hand, it fixes maximum permissible risk of an inefficiency of the investments. We will assume the capital of the investor be equal 1. The problem of optimization of a share portfolio consists in a finding of a vector of share price distribution of papers in a portfolio $x = \{x_i\}, i = \overline{1, N}$ of the investor maximizing the income at the set risk level (obviously, that $\sum_{i=1}^{N} x_i = 1$).

In process of practical application of Markovitz model its lacks were found out:

– The hypothesis about normality profitableness distributions in practice does not prove to be true;
– Stationary of price processes also not always is confirmed in practice;
– At last, the risk of assets is considered as a dispersion (or standard deviation) of the prices of securities from expected value i.e. as decrease in profitableness of securities in relation to expected value, and profitableness increase in relation to an average are estimated absolutely the same.

Though for the proprietor of securities these events are absolutely different.

These weaknesses of Markovitz theory define necessity of use of essentially new approach of definition of an optimum investment portfolio.

Let review the main principles and idea of a method.

The risk of a portfolio is not its volatility, but possibility that expected profitableness of a portfolio will appear below some pre established planned value:

– Correlation of assets in a portfolio is not considered and not accounted;
– Profitableness of each asset is not random, but a fuzzy number. Similarly, restriction on extremely low level of profitableness can be both usual scalar and fuzzy number of any kind. Therefore optimize a portfolio in such statement may mean, in that specific case, the requirement to maximize expected profitableness of a portfolio in a point of time T at the fixed risk level of a portfolio;

– Profitableness of a security on termination of ownership term is expected to be equal r and is in a settlement range. For i -*th* security let's denote:
  - $\bar{r}_i$ – Expected profitableness of *i-th* security;
  - r_{1i} – The lower border of profitableness of *i-th* security;
  - r_{2i} – The upper border of profitableness of *i-th* security;
  - $r_i = \left(r_{1i}, \bar{r}_i, r_{2i}\right)$ – Profitableness of *i-th* security is triangular fuzzy number.

Then profitableness of a portfolio:

$$r = (r_{\min} = \sum_{i=1}^{N} x_i r_{1i}; \bar{r} = \sum_{i=1}^{N} x_i \bar{r}_i; r_{\max} = \sum_{i=1}^{N} x_i r_{2i}) \tag{1}$$

where x_i is a weight of i -*th* asset in portfolio, and

$$\sum_{i=1}^{N} x_i = 1,\ x_i \geq 0,\ i = \overline{1,N} \tag{2}$$

Critical level of profitableness of a portfolio at the moment of T may be fuzzy triangular type number $r^* = \left(r_1^*; \bar{r}^*; r_2^*\right)$.

Direct portfolio optimization problem with triangular membership functions

To define structure of a portfolio which will provide the maximum profitableness at the set risk level, it is required to solve the following problem (3):

$$\{x_{opt}\} = \{x\} \mid r \to \max,\ \ \beta = const \tag{3}$$

where r is profitableness, β is a desired risk, vector's components x satisfy (2).

The most expected value risk degree of a portfolio is defined:

$$\beta = \begin{cases} 0, \ \ if \ \ r^* < r_{\min} \\ R\left(1 + \dfrac{1-\alpha_1}{\alpha_1}\ln(1-a_1)\right), \ \ if \ \ r_{\min} \leq r^* \leq \tilde{r} \\ 1-(1-R)\left(1 + \dfrac{1-\alpha_1}{\alpha_1}\ln(1-a_1)\right), \ \ if \ \ \tilde{r} \leq r^* < \ r_{\max} \\ 1, \ \ if \ \ r^* \geq r_{\max} \end{cases} \tag{4}$$

where

$$R = \begin{cases} \dfrac{r^* - r_{\min}}{r_{\max} - r_{\min}}, & if \quad r^* < r_{\max} \\ 1, \quad if \quad r^* \geq \mathrm{r}_{\max} \end{cases}$$

$$\alpha_1 = \begin{cases} 0, \quad if \quad r^* < r_{\min} \\ \dfrac{r^* - r_{\min}}{\tilde{r} - r_{\min}}, & if \quad r_{\min} \leq r^* < \tilde{r} \\ 1, & if \quad r^* = \tilde{r} \\ \dfrac{r_{\max} - r^*}{r_{\max} - \tilde{r}}, & if \quad \tilde{r} < r^* < r_{\max} \\ 0, & if \quad r^* \geq r_{\max} \end{cases} \tag{5}$$

Having recollected also, that profitableness of a portfolio is:

$$r = (r_{\min} = \sum_{i=1}^{N} x_i r_{1i};\ \tilde{r} = \sum_{i=1}^{N} x_i \tilde{r}_i;\ r_{\min} = \sum x_i r_{2i})$$
,

where $\left(r_{1i}, \overline{r_i}, r_{2i}\right)$ is the profitableness of the i -th security, we obtain the following problem of optimization (6) - (8):

$$\tilde{r} = \sum_{i=1}^{N} x_i \tilde{r}_i \to \max \tag{6}$$

$$\beta = const \tag{7}$$

$$\sum_{i=1}^{N} x_i = 1,\ x_i \geq 0,\ i = \overline{1, N} \tag{8}$$

At a risk level β variation 3 cases are possible. We‘ll consider in detail each of them.

1. $\beta = 0$

From (4) it is evident, that this case is possible when $r^* < \sum_{i=1}^{N} x_i r_{i1} \cdot$.

We obtain the following problem of linear programming:

$$\tilde{r} = \sum_{i=1}^{N} x_i \tilde{r}_i \to \max \tag{9}$$

$$\sum_{i=1}^{N} x_i r_{i1} > r^* \tag{10}$$

$$\sum_{i=1}^{N} x_i = 1,\ x_i \geq 0,\ i = \overline{1, N} \tag{11}$$

Found result of the problem solution (9) - (11) vector $x = \{x_i\},\ i = \overline{1, N}$ is a required structure of an optimum portfolio for the given risk level.

2. $\beta = 1$

From (4) it follows, that this case is possible when $r^* \geq \sum_{i=1}^{N} x_i r_{i2}$.

We obtain the following problem

$$\tilde{r} = \sum_{i=1}^{N} x_i \tilde{r}_i \to \max, \quad \sum_{i=1}^{N} x_i r_{i2} \leq r^*,$$

$$\sum_{i=1}^{N} x_i = 1,\ x_i \geq 0,\ i = \overline{1, N}$$

Found result of the problem decision (9) - (11) vector $x = \{x_i\},\ i = \overline{1, N}$ is a required structure of an optimum portfolio for the given risk level.

3. $0 < \beta < 1$

From (4) it is evident, that this case is possible when $\sum_{i=1}^{N} x_i r_{i1} \leq r^* \leq \sum_{i=1}^{N} x_i \tilde{r}_i$, or when $\sum_{i=1}^{N} x_i \tilde{r}_i \leq r^* \leq \sum_{i=1}^{N} x_i r_{i2}$.

a) Let $\sum_{i=1}^{N} x_i r_{i1} \leq r^* \leq \sum_{i=1}^{N} x_i \tilde{r}_i$ .Then using (4) - (5) problem (6) - (8) is reduced to the following problem of nonlinear programming [Zaychenko, 2006]:

$$\tilde{r} = \sum_{i=1}^{N} x_i \tilde{r}_i \to \max,$$

$$\left(\left(r^* - \sum_{i=1}^{N} x_i r_{i1} \right) + \left(\sum_{i=1}^{N} x_i \tilde{r}_i - r^* \right) \cdot \ln \left(\frac{\sum_{i=1}^{N} x_i \tilde{r}_i - r^*}{\sum_{i=1}^{N} x_i \tilde{r}_i - \sum_{i=1}^{N} x_i r_{i1}} \right) \right) \cdot \tag{12}$$

$$\cdot \frac{1}{\sum_{i=1}^{N} x_i r_{i2} - \sum_{i=1}^{N} x_i r_{i1}} = \beta, \tag{13}$$

$$\sum_{i=1}^{N} x_i r_{i1} \leq r^* \tag{14}$$

$$\sum_{i=1}^{N} x_i \tilde{r}_i > r^* \tag{15}$$

$$\sum_{i=1}^{N} x_i = 1,\ x_i \ge 0,\ i = \overline{1,N} \quad (16)$$

б) Let $\sum_{i=1}^{N} x_i \tilde{r}_i \le r^* \le \sum_{i=1}^{N} x_i r_{i2}$. Then the problem (6) - (8) is reduced to the following problem of nonlinear programming:

$$\tilde{r} = \sum_{i=1}^{N} x_i \tilde{r}_i \to \max \quad (17)$$

$$\left(\left(r^* - \sum_{i=1}^{N} x_i r_{i1} \right) - \left(r^* - \sum_{i=1}^{N} x_i \tilde{r}_i \right) \cdot \ln \left(\frac{r^* - \sum_{i=1}^{N} x_i \tilde{r}_i}{\sum_{i=1}^{N} x_i r_{i2} - \sum_{i=1}^{N} x_i r_{i1}} \right) \right) \cdot \frac{1}{\sum_{i=1}^{N} x_i r_{i2} - \sum_{i=1}^{N} x_i r_{i1}} = \beta \quad (18)$$

$$\sum_{i=1}^{N} x_i r_{i2} > r^* \quad (19)$$

$$\sum_{i=1}^{N} x_i \tilde{r}_i \le r^* \quad (20)$$

$$\sum_{i=1}^{N} x_i = 1,\ x_i \ge 0,\ i = \overline{1,N} \quad (21)$$

The R-algorithm of minimization of not differentiated functions is applied to the decision of problems (12) - (16) and (17) - (21). Let both problems: (12) - (16) and (17) - (21) solvable. Then to the structure of a required optimum portfolio will correspond a vector – $x = \{x_i\},\ i = \overline{1,N}$ the decision of that problem (12) - (16), (17) - (21) the criterion function value of which will be greater.

The dual optimization problem

It is necessary to determine the structure of the portfolio, which will provide a minimum level of risk for a given level of the portfolio profitability.

We obtain the following optimization problem:

$$\min \beta(x),\ \tilde{r} = \sum_{i=1}^{N} x_i \tilde{r}_i \ge r_{зад} = \mathrm{r}^*, \quad (22)$$
$$\sum_{i=1}^{N} x_i = 1,\ x_i \ge 0,\ i = \overline{1,N},$$

where r and β is determined by the used membership function.

Consider optimization problem with the triangular MF.

It is necessary to solve optimization problem (22) where $\beta(x)$ is determined from the formula (4) and (5).

Multicriteria optimization problem

Now consider multicriteria fuzzy portfolio optimization problem in which portfolio profitableness should be maximized and risk should be minimized.

In order to find the structure of corresponding fuzzy portfolio the following problem is to be solved:

$$\{x_{opt}\} = \{x\} \quad | \quad r \to \max, \quad \beta \to \min, \tag{23}$$

where r and β are determined by formulas (4) and (5) and vector X components satisfy (8).

For simplifying problem solution transfer it to single criterion. Normalize the value of profitableness as follows:

$$\tilde{r}_H = \frac{r_{\max} - \tilde{r}}{r_{\max} - r_{\min}}, \quad \tilde{r}_H \in [0;1], \tag{24}$$

Using formulas (23) and (24), we obtain the optimization problem in such form:

$$\{w_1 \tilde{r}_H + w_2 \beta(x)\} \to \min$$
$$w_1 \ge 0, \; w_2 \ge 0, \; w_1 \ne w_2, \; w_1 + w_2 = 1$$
$$\sum_{i=1}^{N} x_i = 1, \; x_i \ge 0, \; i = \overline{1, N}$$

Optimization problem during the time period

In this case we must define the structure of the portfolio, which provides the maximum average return for a given level of risk. So we calculate the profitableness from (3) as:

$$\tilde{r} = \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N} x_{ti} \tilde{r}_{ti}, \tag{25}$$

where $\tilde{r}_{ti}$ - the expected profitableness of *i*-th security in time unit t. T - the length of time period. We should find an optimal portfolio from the (3), using (4), (5), (25).

Analysis of the results

The profitableness values of leading companies in the period from 01.12.2014 to 10.04.2015 were used as the input data. The companies: Google Inc (GOOGL), Walt Disney Co (DIS), The Coca Cola Co (KO), Kimberly Clark Corp (KMB), Seagate Technology PLC (STX), Tesla Motors Inc (TSLA). The corresponding data is presented in the Table 1.

Table 1. The profitableness

Companies / Dates	**GOOGL**	**DIS**	**KO**	**KMB**	**STX**	**TSLA**
05.12.2014	-0,0214	0,0114	-0,0229	-0,0161	-0,0192	-0,0418
12.12.2014	-0,0174	-0,0246	-0,0517	-0,0126	-0,0312	-0,0143
12.12.2014	0,0081	0,0219	0,0340	0,0311	0,0847	0,0389
26.12.2014	0,0173	0,0088	0,0144	0,0157	-0,0067	-0,0284
02.01.2015	-0,0144	-0,0183	-0,0168	-0,0161	-0,0235	-0,0163
09.01.2015	-0,0361	0,0202	0,0211	0,0091	0,0185	-0,0452
16.01.2015	0,0270	0,0076	-0,0026	0,0163	-0,0207	0,0488
23.01.2015	0,0628	-0,0002	0,0035	-0,0500	0,0136	-0,0143
30.01.2015	0,0015	-0,0422	-0,0426	-0,0172	-0,0444	0,0304
06.02.2015	0,0032	0,1098	-0,0034	-0,0096	0,0370	-0,0630
13.02.2015	0,0413	0,0240	0,0184	0,0257	0,0238	0,0624
20.02.2015	-0,0059	0,0041	0,0038	-0,0028	0,0168	-0,0173
27.02.2015	0,0516	-0,0087	0,0346	-0,0120	-0,0140	-0,0186
06.03.2015	-0,0037	-0,0195	-0,0389	-0,0325	-0,0757	-0,0174
13.03.2015	-0,0368	0,0112	-0,0358	-0,0115	-0,0565	-0,0115
20.03.2015	0,0059	0,0099	0,0089	0,0193	0,0305	0,0124
27.03.2015	**-0,0138**	**-0,0253**	**-0,0133**	**-0,0240**	**-0,0553**	**-0,0733**
02.04.2015	**-0,0353**	**-0,0011**	**0,0042**	**-0,0067**	**-0,0119**	**0,0023**
10.04.2015	**0,0084**	**0,0125**	**-0,0070**	**-0,0056**	**0,0538**	**0,0384**

For getting an input data for the optimization system we used the Fuzzy GMDH method with triangular membership functions, linear partial descriptions, training sample of 70% size. The profitableness values were forecasted for each of 3 weeks (Table 2).

Table 2. Forecasted profitableness

Companies	Profitableness				MAPE test sample	MSE test sample
	Real value	Low bound	Forecasted value	Upper bound		
27.03.2015						
GOOGL	-0,0138	-0,0408	-0,0116	0,0176	1,5116	0,0265
DIS	-0,0253	-0,0130	0,0087	0,0304	2,0415	0,0412
KO	-0,0133	-0,0629	-0,0160	0,0309	1,0619	0,0131
KMB	-0,0240	-0,1520	-0,0303	0,0914	2,3122	0,0136
STX	-0,0553	-0,0618	-0,0344	-0,0070	2,5215	0,0462
TSLA	-0,0733	-0,1919	-0,0391	0,1137	1,459	0,0078
02.04.2015						
GOOGL	-0,0353	-0,0672	-0,0328	0,0016	2,0537	0,0226
DIS	-0,0011	0,0082	0,0299	0,0516	1,7141	0,0371
KO	0,0042	-0,0418	0,0021	0,0460	1,9743	0,0159
KMB	-0,0067	-0,1342	-0,0125	0,1092	1,7452	0,0245
STX	-0,0119	-0,0164	0,0090	0,0344	1,8243	0,0172
TSLA	0,0023	-0,0999	0,0349	0,1697	2,101	0,0241

10.04.2015						
GOOGL	0,0084	-0,0215	0,0109	0,0433	1,439	0,0162
DIS	0,0125	0,0308	0,0425	0,0542	3,103	0,0215
KO	-0,007	-0,0506	-0,0087	0,0332	1,037	0,0107
KMB	-0,0056	-0,1136	-0,0119	0,0898	2,014	0,0135
STX	0,0538	0,0483	0,0747	0,1011	2,855	0,0178
TSLA	0,0384	0,0362	0,069	0,1018	2,014	0,0194

In this way the portfolio optimization system stops to be dependent on factor of expert subjectivity. Besides, we can get data for this method automatically, without expert's estimates.

Let the critical profitableness level set by 7,5 %. Varying the risk level we obtain the following results for triangular MF (10.04.2015). The results are presented in the Tables 3, 4 and Figure 1.

Table 3. Direct problem - distribution of components of the optimal portfolio with critical level r *=7,5 %

GOOGL	DIS	KO	KMB	STX	TSLA
0,00065	0,00025	0,00137	0,00423	0,99017	0,00333
0,00159	0,00157	0,00066	0,00323	0,9923	0,00065
0,00435	0,00307	0,00058	0,00271	0,98533	0,00396
0,00018	0,00008	0,00082	0,00291	0,99363	0,00238
0,0049	0,00394	0,00071	0,00244	0,97964	0,00837
0,00287	0,00525	0,00296	0,00506	0,97678	0,00708
0,0033	0,00187	0,00279	0,00301	0,9707	0,01833

0,00188	0,00232	0,00153	0,0125	0,97369	0,00808
0,00439	0,00297	0,00238	0,0152	0,94177	0,03329
0,00314	0,00317	0,0036	0,02553	0,93466	0,0299
0,00339	0,00579	0,00405	0,03727	0,93124	0,01826

Table 4. Direct problem - parameters of the optimal portfolio with critical level r *=7,5 %

Low bound	Expected profitableness	Upper bound	Risk level
4,732	**7,412**	**10,093**	**0,25**
4,756	7,421	10,085	0,3
4,74	7,402	10,064	0,35
4,77	7,435	10,1	0,4
4,732	7,394	10,056	0,45
4,681	7,362	10,043	0,5
4,705	7,383	10,061	0,55
4,586	7,325	10,064	0,6
4,484	7,262	10,04	0,65
4,317	7,172	10,026	0,7
4,131	7,063	9,995	0,75

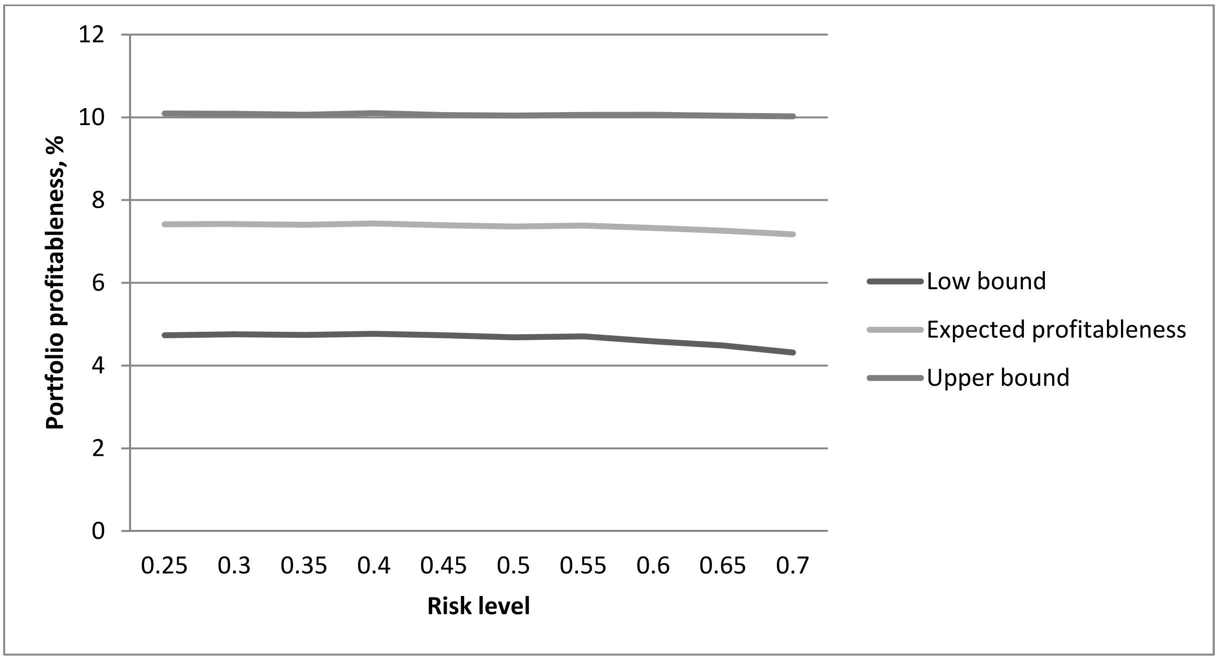


Figure 1. Dependence of expected portfolio profitableness on risk level

As we can see on Figure 1 the dependence profitableness - risk has a descending type, the greater risk the lesser is profitableness opposite to classical probabilistic methods. It may be explained so that at fuzzy approach by risk is meant the situation when the expected profitableness happens to be less than the given criteria level. When the expected profitableness decreases, the risk grows.

The profitableness of the real portfolio is 5,34 %.This value falls in results calculated corridor of profitableness [4,732; 7,412; 10,093], indicating the high quality of the forecast. The main portfolio portion in this case goes to company Seagate Technology PLC that can be explained by the high level of its profitableness in comparison with other companies.

Now consider in Tables 5, 6 and Figure 2, the results obtained by solving multicriteria problem.

Table 5. Multicriteria problem - distribution of components of the optimal portfolio with critical level r *=7,5 %

GOOGL	DIS	KO	KMB	STX	TSLA
0,01069	0,02289	0,00316	0,00117	0,92946	0,03263
0,0121	0,02234	0,0058	0,00414	0,92512	0,0305
0,01265	0,02131	0,00733	0,00596	0,92451	0,02824

0,01418	0,02124	0,00982	0,00871	0,91908	0,02697
0,01478	0,02049	0,01124	0,01038	0,918	0,02511
0,0154	0,01989	0,01265	0,01199	0,91656	0,02351
0,01603	0,01929	0,01402	0,01355	0,91518	0,02193
0,01668	0,0188	0,01537	0,01505	0,91356	0,02054
0,01733	0,01838	0,01669	0,01655	0,91182	0,01923
0,01069	0,02289	0,00316	0,00117	0,92946	0,03263
0,0121	0,02234	0,0058	0,00414	0,92512	0,0305

Table 6. Multicriteria problem - parameters of the optimal portfolio with critical level r *=7,5 %

Low bound	Expected profitableness	Upper bound	Risk level	w1
4,626	**7,273**	**9,92**	**0,65253**	**0,1**
4,545	7,219	9,893	0,67539	0,2
4,501	7,192	9,882	0,69878	0,3
4,423	7,138	9,854	0,71007	0,4
4,381	7,112	9,842	0,72732	0,5
4,34	7,085	9,83	0,74142	0,6
4,3	7,059	9,818	0,75483	0,7
4,26	7,033	9,806	0,76619	0,8
4,221	7,007	9,793	0,77641	0,9

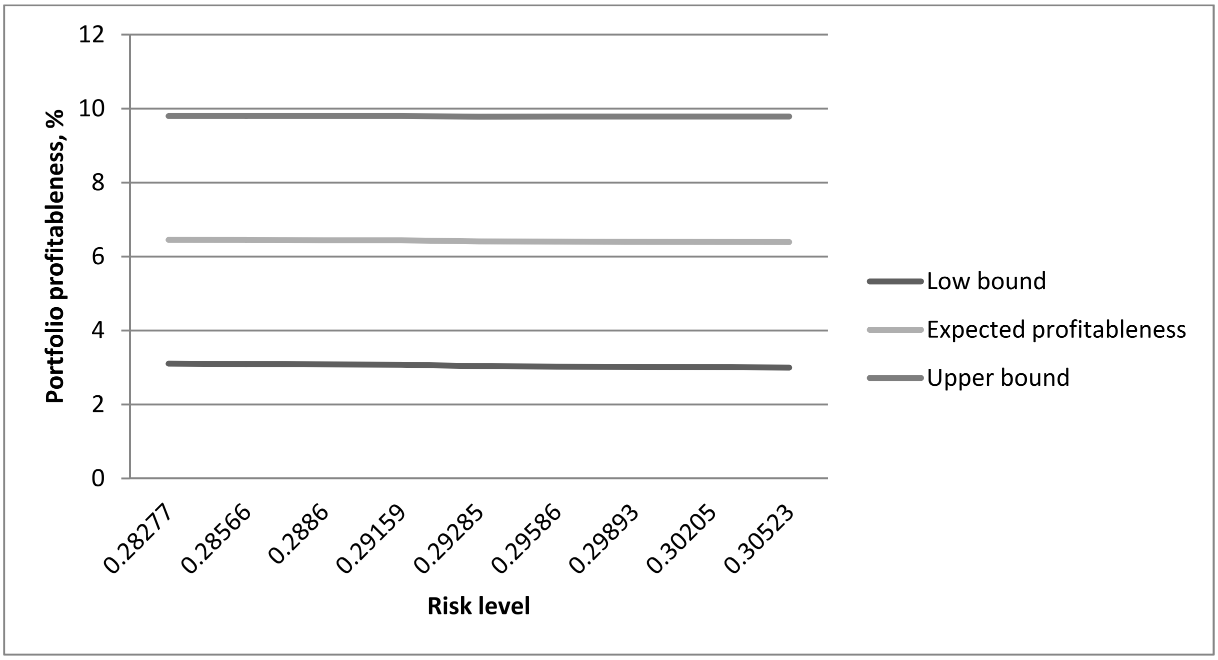


Figure 2. Dependence of expected portfolio profitableness on risk level

The profitableness of the real portfolio is 5,16 %. This value falls in results calculated corridor of profitableness [4,626; 7,273; 9,92]. The dependence profitableness - risk also has descending type.

Let’s consider an optimization portfolio at the time interval of 3 weeks. We have obtained profitableness values for all weeks by using the Fuzzy GMDH [Zaychenko, 2000] (Tables 7, 8 and Figure 3).

Table 7. Optimization problem during the period - distribution of components of the optimal portfolio for triangular MF with critical level r *=7,5 %

GOOGL	DIS	KO	KMB	STX	TSLA
0,01964	0,92474	0,02535	0,00102	0,02878	0,00047
0,01965	**0,92309**	**0,02536**	**0,00287**	**0,02876**	**0,00027**
0,02066	0,92203	0,02637	0,00066	0,02975	0,00053
0,01968	0,92274	0,02538	0,00069	0,02873	0,00278
0,01969	0,92329	0,02539	0,00206	0,02871	0,00086

0,0197	0,92068	0,0254	0,00208	0,02869	0,00345
0,01972	0,92355	0,02541	0,00231	0,02867	0,00034
0,01973	0,9228	0,02542	0,00223	0,02865	0,00117
0,02075	0,9212	0,02643	0,00124	0,02962	0,00076
0,01976	0,92152	0,02544	0,00205	0,0286	0,00263

Table 8. Optimization problem during the period - parameters of the optimal portfolio for triangular MF with critical level r *=7,5 %

Low bound	Expected profitableness	Upper bound	Risk level
4,732	7,412	10,093	0,25
4,756	**7,421**	**10,085**	**0,3**
4,74	7,402	10,064	0,35
4,77	7,435	10,1	0,4
4,732	7,394	10,056	0,45
4,681	7,362	10,043	0,5
4,705	7,383	10,061	0,55
4,586	7,325	10,064	0,6
4,484	7,262	10,04	0,65
4,317	7,172	10,026	0,7

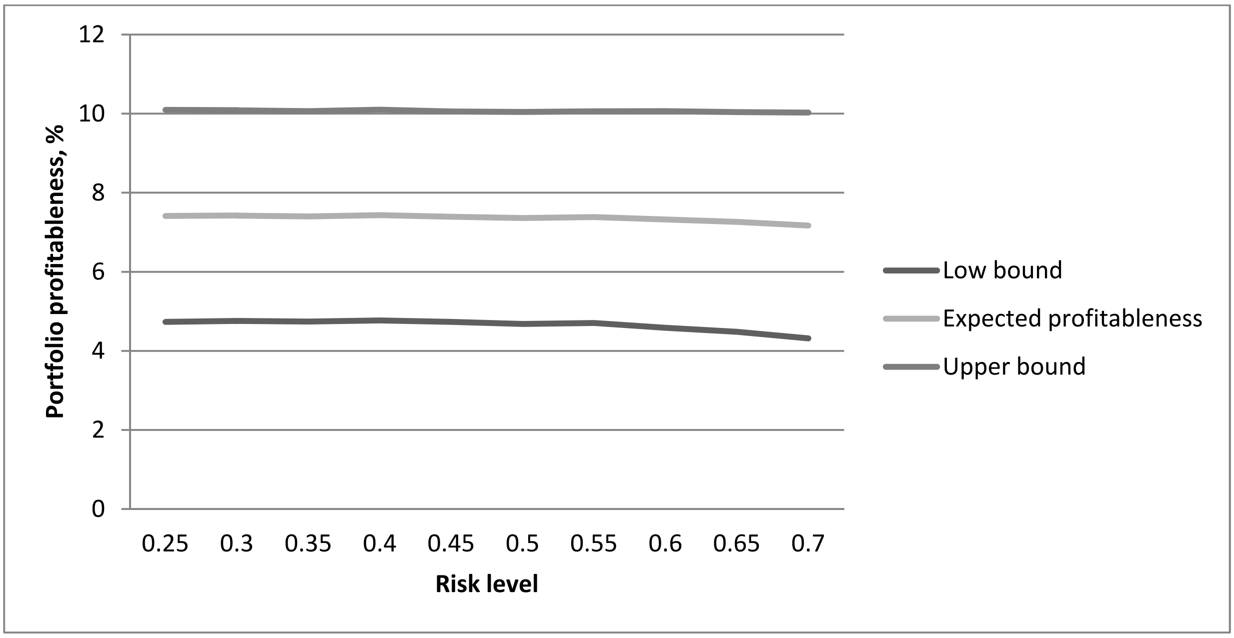


Figure 3. Dependence of expected portfolio profitableness on risk level

Let's consider the results obtained by solving the dual problem using triangular MF. In this case, the investor sets the rate of return, and the problem is to minimize the risk. The main portfolio portion in this case goes to company Walt Disney Co that can be explained by the high level of its average profitableness in comparison with other companies.

The optimal portfolio is listed in Tables 9, 10 and the dependence of the risk level on a given critical return is presented in the Figure 4.

Table 9. Dual problem. Distribution of components of the optimal portfolio

GOOGL	DIS	KO	KMB	STX	TSLA
0,01756	0,01918	0,01664	0,01503	0,91202	0,01956
0,01752	0,02002	0,01609	0,01389	0,91166	0,02082
0,01666	0,02021	0,01459	0,01194	0,91494	0,02166
0,01606	0,02093	0,01318	0,01032	0,91606	0,02345
0,01404	0,02119	0,00973	0,00691	0,92231	0,02582
0,00768	0,02313	0,00185	0,00015	0,93191	0,03528
0,00951	0,02323	0,0009	0,0006	0,93027	0,0355

Table 10. Dual problem. Parameters of the optimal portfolio

Low bound	Expected profitableness	Upper bound	Risk level	Critical rate of return
0,07016	0,04242	0,0979	0,01821	5
0,07027	0,04263	0,09791	0,05381	5,5
0,07061	0,04315	0,09808	0,11501	6
0,07088	0,04356	0,0982	0,21157	6,5
0,07157	0,04456	0,09857	0,37038	7
0,0731	0,04673	0,09947	0,6253	7,5
0,07302	0,04661	0,09942	0,87555	8

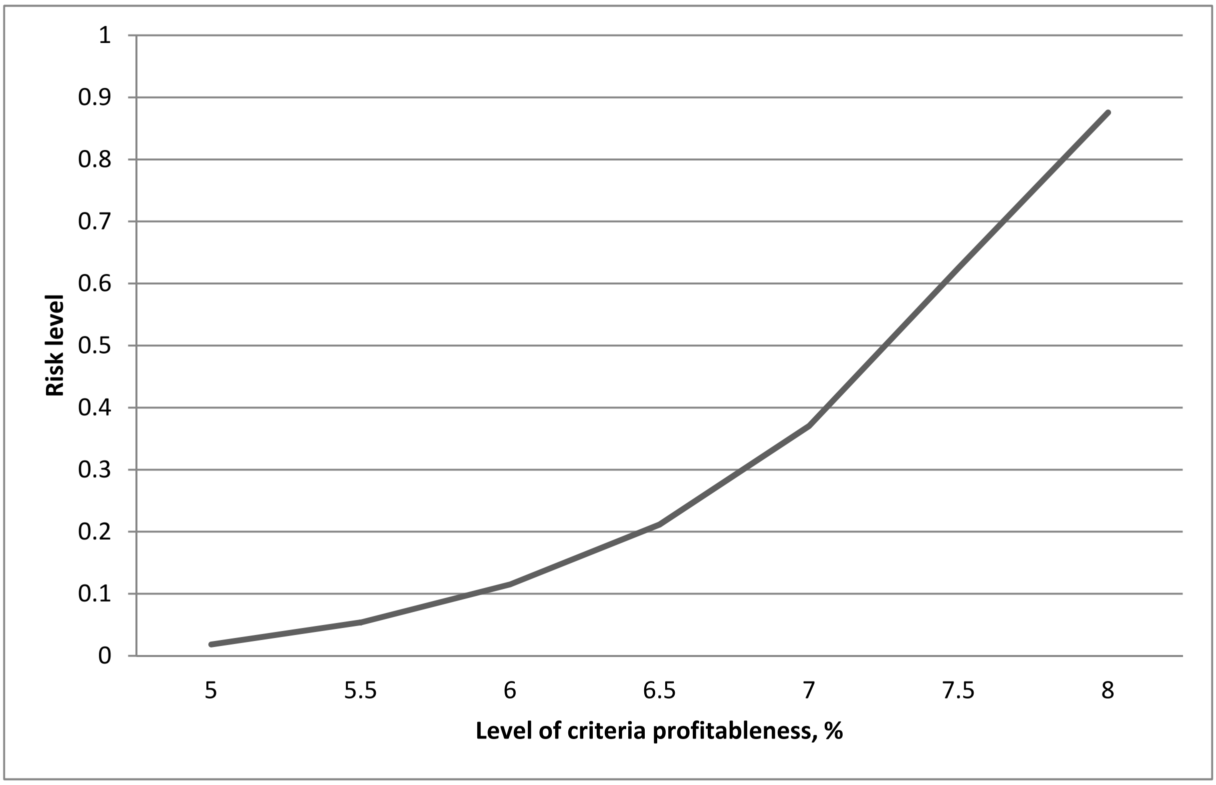


Figure 4. Dependence of the risk level on a given critical return

From these results we can see that the dependence risk - given critical level of profitability takes a growing character, because at the growth of the critical profitability the probability that the expected return will be lower than a given critical value also increases.

Conclusion

In this work the research in the field of portfolio management was carried out. Fuzzy sets theory was used as a tool for getting an optimal portfolio. As a result of research the mathematical model based on the fuzzy-set approach for a finding of structure of the optimal investment portfolio has been obtained. On the basis of the theory of fuzzy sets the algorithm of optimization of a share portfolio has been developed. As a result of research the following conclusions were made:

– The dependence profitableness - risk has a descending type, the greater risk the lesser is profitableness opposite from classical probabilistic methods. It may be explained so that at fuzzy approach by risk is meant the situation when the expected profitableness happens to be less than the given criteria level. When the expected profitableness decreases, the risk grows;
– Portfolios during the time period and at the end of period have different structure and characteristics, that can be explained by the variations of average profitableness;
– For improving the accuracy of the suggested fuzzy portfolio model, the fuzzy GMDH method was applied for profitableness forecasting. The experimental investigations have proved its high efficiency.
– The dependence risk - given critical level of profitability has a growing character in the dual task, because with the growth of the critical profitability increases the probability that the expected return will be lower than a given critical value.

Thus, we have developed a system that not only automates the search of the optimal portfolio, but also provides a flexible and effective management of portfolio investments.

Bibliography

[Burenin, 1998] Буренин А. Н. Рынок ценных бумаг и производных финансовых инструментов [Текст]: учеб. пособие / Буренин Алексей Николаевич, М.: 1 Федеративная Книготорговая Компания,1998, 352 с.

[Nedosekin, 2002] Недосекин А.О., "Нечетко-множественный анализ риска фондовых инвестиций" [электронный ресурс], Режим доступа: logic-bratsk.ru/radio/fuzzy/nedosek/book23.pdf

[Zadeh, 1999] Zadeh L.A. "Fuzzy sets as a basis for a theory of possibility", [електронний ресурс], Режим доступу: http://portal.acm.org/citation.cfm?id=310820

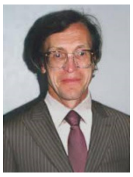
[Zaychenko & Maliheh, 2008] Зайченко Ю.П., Малихех Есфандиярфард, "Анализ инвестиционного портфеля для различных видов функций принадлежности" //Системні дослідження та інформаційні технології, №2, 2008, с. 59-76.

[Zaychenko, 2000] Зайченко Ю.П. Нечіткий метод групового врахування аргументів та його застосування в задачах прогнозування макроекономічних показників[Текст]/ Ю.П. Зайченко, О.Г. Кебкал, В.Ф. Крачковський// Наукові вісті НТУУ "КПІ", №2, 2000, с. 18-26.

[Zaychenko, 2006] Зайченко Ю.П. "Дослідження операцій [Текст]: підручник / Юрій Петрович Зайченко, К.: Слово, 2006, 816 с.

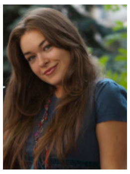
[Zaychenko, 2008] Зайченко Ю.П., “Нечеткие модели и методы в интеллектуальных системах”, [Текст]: учеб. пособие для студентов высших учеб. заведений / Юрий Петрович Зайченко, К.: Слово, 2008, 341с.

Authors' Information



Yuri Zaychenko *– Proffesor, doctor of technical sciences, Institute for applied system analysis, NTUU “KPI”, 03056, Ukraine, Kyiv, Peremogi pr. 37, Corpus 35; e-mail: baskervil@voliacable.com*

Major Fields of Scientific Research: Information systems, Fuzzy logic, Decision making theory, Fuzzy sets theory, Investment portfolio optimization, Modeling and optimization of computer networks



Inna Sydoruk *– Postgraduate, Institute for applied system analysis, NTUU “KPI”, 03056, Ukraine, Kyiv, Peremogi pr. 37, Corpus 35; e-mail: isidoruk@ukr.net*

Major Fields of Scientific Research: Information systems, Fuzzy sets theory, Investment portfolio optimization, Decision making theory

REVIEW OF THE SOFTWARE SOLUTIONS OF LOGISTIC TASKS

Vitalii Lytvynov, Alina Posadska

Abstract: *Description of software solutions for logistic tasks and their comparative analysis are considered in the article. The selection of the functionality of Logistics Information System, which is designed to automate the planning of the transport delivery of freights, is presented. The logistical tasks considering features of transportation of agricultural freights, which should solve the Logistics Information System, are analyzed.*

Keywords: *logistics information system, automation of delivery routes, temporary infrastructure, GIS – technology, ERP-systems.*

ACM Classification Keywords: *D.2.0 Genera - Software Engineering.*

1. Introduction

Formulation of the problem. Logistics is the science and practical methods of management of material, financial and information flows.

Based on the related work in the field [Гончаров, 2012], the conclusion can be made that the methods and tools for solving logistic tasks are becoming more demanded for such distributed system as agro-industrial complex.

The necessity of usage of logistics systems in the agricultural sector is associated with the fact that agricultural industries characterized by a lack of territorial localization of production processes. Enterprises processing agricultural products of industry, usually spatially removed from raw material sources, which necessitates the physical movement of the material flow in both time and space.

The main tasks of logistics are overcoming the distance during the movement of freights, time tracking and organization of effective service supply. At the same time, transportation, storage, maintenance of orders for products are the activities that have great importance for the effectiveness of any enterprise.

Analysis of recent research and publications. Investigations in the area of logistics are designed to solve economic problems and tasks related to the process of movement of goods, funds and information.

In case of the development of information technologies and the usage of logistic software solutions in agricultural firms, there is a problem of integration of logistics information systems because of the specific features of enterprises work, which is caused not only by the geographical distribution, but also by the strong impact of environmental factors on the transport system

Territorial distribution of transport systems makes them an ideal object for automation using the means of geographic information systems (GIS).

GIS is an optimal platform for complex solutions in the area of transport, since the spatial component is the natural basis for the integration of task of transport infrastructure management, solutions of computing tasks, tasks of operational management, navigation, etc [Разгуляев, 2010]. However, these solutions have specific nature because of ignoring the influence of external factors. Real complex solutions in this area have not offered in Ukraine yet.

Such situation is typical not only for Ukraine. Analysis of the works of domestic and foreign authors shows that there is no holistic logistics concept of the organization of transport operations in the agricultural enterprises by now [Костышева, 2013].

Naturally, current absence leads to a lack of tools for automation of logistics process management.

The purpose of current work is a comparative analysis of existing software products for the solution of logistical problems applied to agricultural enterprises.

2. The main tasks of Logistics Information System

The agricultural enterprise is complicated distributed object, which consists of many other different control objects [Бальченко, 2013]. Their functioning depends on impact of different factors (weather conditions, social environment, quality of resources used, quality of work performance, etc.) [Lytvynov, 2014].

The object of automation are the processes of planning optimal routes of delivery of commercial products to processing end-points, intermediate temporary storage points, warehouse managing, and controlling efficiency of these processes implementation, minimizing the products, financial and material losses during the implementation of logistics processes in agricultural enterprise, minimizing of possible risks.

At the stage of tasks setting it is necessary to consider the features of agricultural freights transportations: seasonality in the harvest that leads to fluctuations in freight turnover and traffic volume; short terms of harvesting, which require hard work of auto transport; unevenness of crops maturation in different climatic and soil regions of the country; fluctuation crop capacity during drought and other adverse weather conditions; heavy traffic work conditions of rolling stock, particularly in spring and autumn; low volume weight of agricultural goods, which does not allow full use of the load capacity of rolling stock.

Logistics Information System (LIS) is intended for complex informational, analytical, expert, predictive, optimizational ensuring of logistic processes of agricultural enterprise for the following tasks:

1. Planning and automation of route:

- planning of optimal delivery routes - automated control system (A&C) of the transport should automatically calculate routes and suggest the best option or typical delivery route that takes into account the type of vehicle, the degree of loading transport, time windows, randomness in time (time of day, weekend/weekday, season), and other criteria which can be adjusted according to the specific requirements of the company;
- development of operational planning schedules of traffic with a binding to work schedule of processing capacities of enterprise, state of transport units, meteorological forecasts;
- planning using different types of transport (tractor side vehicle, truck) - transportation planning, which takes into account the type of freight and its transportation features, will protect the company from force majeure and empty runs of vehicles - it will reduce transport costs;
- accounting of real road network (forbidden turns, one-way traffic, temporary roads), and paving roads that are not present on the map;
- accounting of different conditions during route planning: distances between points, cost of using cars, characteristics of freights and transport, time of loading/unloading and execution of documents, weather conditions and other;
- accounting of restrictions of delivery and features of freights: correcting of terms and conditions of delivery perishable freights, determination of losses during transport.

2. Planning and management of maintenance and repair of rolling stock.

3. Forecasting, distribution and planning of material resources:

- inventory management in storages of auto transport park and organization of storing materials at storages;
- accounting of deadlines storage of resources;
- planning of the usage of additional resources.

4. Planning in stationary conditions and in conditions of infrastructure changes. Infrastructure changes imply a dynamic change in the state of transport, fields, roads, and the use of temporary infrastructure (temporary storage).

- Developing of programs for economy of material resources, transport usage and control over their implementation: monitoring the usage of vehicles, fuel, kilometrage, transportation costs, travel time, number of involved vehicles, waiting time.

These tasks must be solved in real time and in a binding to current state of transport system.

These solutions require hardware-software tools for automatic collection of primary information and also software package.

3. Software solutions of logistic tasks

The analysis of the software products presented at IT-market has been conducted.

All logistics software can be separated into two categories:

- logistics software based on GIS – technology;
- logistics software based on accounting-optimization tasks.

GIS systems are more appropriate for solving transport logistic tasks. There is classification of their possibilities:

1. electronic map with the means of routing and road navigation, which also includes a universal reference system;
2. software products for monitoring the location and status of mobile objects (transport, freights, sales representatives), which are intended for solving such tasks as tracking the location and status of transport and freight; monitoring of implementation of schedule and route (deviation from plan);
3. software products for automatic planning of complex delivery with automatic controlling of parameters and the possibility of manual correction of calculated runs;
4. software products for complex automation of business processes of transport enterprise management.

Currently, the most demanded software products in IT-market are electronic maps with automatic making of routes. But companies that solve tasks of transport logistics, prefer software products of classes 2, 3 and 4 of the above classification [Бочкарев, 2013].

Usually, these modules are embedded into such products as MAPINFO, ARC/INFO, Optimum GIS, etc [Сергеев, 2008].

OPTIMUM GIS

Geoinformation system OPTIMUM GIS (developed by Group CDC, Russia) consists of several modules, which are designed for automation of transport logistics (planning of delivery routes of freights), satellite GPS/GLONASS monitoring of transport, forming of routes and GPS/GLONASS monitoring of field staff, all mobile workers (sales representatives, merchandisers, operational maintenance teams, service engineers, bank agents, etc.) [CDC official website, 2014].

OPTIMUM GIS modules for automation of transport logistics:

- Planning of routes of freights delivery (module «Delivery»).
- GPS/GLONASS monitoring of transport (module «Monitoring»).
- Automation of the forwarder (module «Forwarder»).

Related solutions:

- OPTOKEYS. Mobile solution for printing and storage of documents.
- GPS/GLONASS terminals and trackers.

Modules «Monitoring» and «Delivery», which are included in the system Optimum, are designed for automation of transport logistics processes and tracking of transport movement. Monitoring of implementation of current route is based on a predetermined move schedule on the key points of the route (control point, point of parking, refueling point, etc.). Current location and route are determined by coordinates of transport GPS-tracker, which should be transferred to server part of the system OPTIMUM by using GPRS/SMS through a network of mobile operators.

Optimum GIS module «Monitoring» allows reducing the costs from unjustified downtime of vehicles and route deviations, and provides the to-date information about the location and movement of vehicles on electronic map. The program allows controlling indexes of any sensors, which are installed in vehicle (fuel consumption, opening doors, glass broking, etc.).

Optimum GIS module «Delivery» automatically calculates the optimum loading of auto park, based on existing orders, the route duration, which accounts spent time for discharge and total kilometrage, and returning to storage to fill up. The program allows optimizing of route planning based on the density distribution of points, keeps accounting of «delivery windows», volume of the bulk body, type and class of transport, load capacity and prioritization of points. Module «Delivery» prints set of necessary documents: routing waybill, loading invoice, report forwarder, a paper copy of waybill, bill of lading, invoice, delivery note and other.

The license cost is $ 1,700.

Disadvantages: the program does not have its own protection.

Conclusion: the product is used in agroholdings to improve mobile commerce – it allows to create optimal routes of visiting the sale points, to define certain working scenarios, to monitor their implementation, to control location of mobile workers using gps-monitoring, to reduce the cost of fuel due to the notes of accurate kilometrage in itinerary list, optimizing the routes and control fuel consumption. It can solve only the part of assigned tasks.

Chronomap

Chronomap (developed by the company Mapping Information Systems Corporation, USA) - is a software module that is designed for simulation of transport networks, solution of various tasks of transport and geomarketing. It allows to perform calculations, simulate and visualize solutions, using the most complete and accurate database of digital road maps, increasing the efficiency of the company [MapInfo official website, 2014]. ChronoMap is used by trading networks, banks, service organizations, government agencies, and consulting firms around the world, which need to solve logistical tasks. Modeling of transport network: the program works with any data of the road network in vector format MapInfo Professional. Feature of Chronomap is the ability to define rules for the construction of transport networks that allows bringing the mathematical model (graph) to the actual road conditions as close, as it is possible. The result will be more correct if the more rules will be defined for the transport network. They are including:

- Identify the characteristics of high-speed network segment – road class which has assigned speed characteristics.
- Setting different speed characteristics for the segment in different directions (for two-way streets).
- Setting areas with one-way traffic.
- Reducing the speed characteristics of the network at peak times or in bad weather.
- Prohibition of passage.
- Temporary closing of segment without rebuilding of network graph (for example, repair).
- Prohibition of turns.
- Prohibition on passage of certain kinds of transport (dimensions, tonnage, type, etc.).
- Library of different vehicles.

Also, there is a possibility of selection of the vehicle (car, bike, truck, tractor, etc.) to navigate the route, which is very important. The parameters of each vehicle may be modified: speed, reduction of speed in the rush hour, etc. It is also possible to change the physical characteristics of the vehicles (weight, dimensions, etc.). It allows creating and comparing of different scenarios of availability of certain point with estimating efficiency of each option.

ChronoMap has a function of finding the shortest way, calculation of routes from the starting point via intermediate points to a point of arrival. Optimizing functions allow to define the best route in terms of time, distance or transportation costs, and to optimize the sequence of «stops» (Traveling Salesman Problem).

The license cost is $ 9500.

Advantages: relative simplicity in mastering of the product; simple and clear licensing policy - buy one package and that is enough; the presence of all major functions; easy to learn and convenience of vector cartographic editor; the familiar MS Office-like interface; very high degree of stability in the work.

Disadvantages: weak mathematical support of processing spatial (vector, raster) and attribute data. However, problem is partially solved with help of various extensions, such as Vertical Mapper; interface of georeferencing raster maps is not very user friendly; typification of spatial data is not too strict; selection of built-in standard rules of validation topology is not too large.

Conclusion: it is particularly suitable for solving the problems of transport logistics for our LIS, but considering the drawbacks, it is better to choose another GIS system, for example, ArcGIS.

ArcGIS

ArcGIS (developed by Environmental Systems Research Institute (ESRI), California, USA) is a GIS-system that is the basis for the software products that are designed for complex vehicle fleet management of company or group of companies [ArcGIS official website, 2014]:

- receives data from A&C of enterprise;
- calculates optimal routes for transport via one of two main criteria: a minimum cost or a minimum of time;
- provides the results on any background map substrate (raster maps, vector maps, satellite images, aerial photos, WEB-services); in the form of itinerary sheets with a detailed description of route and stopping points, calculation of run and fuel consumption; in the form of tracks for use in mobile navigators;
- receives and processes data of GPS-sensors on vehicles in real-time;
- controls the route and traffic schedule, estimated time of arrival at destination, location in certain areas; to promptly notify the system operator and other interested persons about the situation using client application by email, SMS, etc.

Transport logistics: any complexity and detail of the road network model for calculations - from schematic to maximum details with according the type of road surface, traffic rules, restrictions on the dimensions of the vehicle, traffic intensity, etc. Monitoring in real-time mode: speed, idling/stop, territories, dangerous areas, arrival/departure, arrival time, observance of route. It is possible using additional modules Route Planner, ArcGIS Spatial Analyst, ArcGIS 3D Analyst, ArcGIS Geostatistical Analyst, ArcGIS Network Analyst and other.

The license cost is from $2100, but it is necessary to buy modules additionally.

Advantages: product has strict data topology; integrity control and data topology using geodatabase; developed apparatus of work with datums, coordinate systems and geographical projections; a well-developed mathematical apparatus of processing of spatial and attribute data.

Disadvantages appear with the introduction of monitoring projects, where update speed is critical:

- licensing policy is confusing and not always clear from the first time. At first, it is unknown what right tool is absent in the purchased license to solve task. Only trained specialist will be able to understand the types of licenses and functionality. The product is not very democratic;
- editor is a little overloaded for mastering – it is not necessary to operator to have a university education;

- bad testing of first assemblies versions (unfortunately, error of program happens in the first releases);
- technical support: there is a huge amount of documentation, a lot of articles. But almost all documentation in English, which is critical for the local user.

Conclusion**:** product partially covers the tasks of transport logistics in agriculture. But the rest of the field was not affected - the warehouse and distribution logistics, inventory, procurement. But it doesn't work with storage and distribution logistics, inventory, procurement. The modifying and integrating with other systems must be performed.

Also there are independent software products that implement separate logistic functions - **modules of large ERP-systems.**

Minus: they require the acquisition and implementation of expensive «basic» system (often - primitive logic, using a few simple methods).

Plus: working in one information field - problems of compatibility of data are minimized.

Microsoft Dynamics AX (Axapta)

Axapta is one of the solutions for enterprise management, which delivers Microsoft Dynamics Division of Microsoft Corporation. Multi-functional ERP is a system of enterprise resource management for medium and large companies. It covers all areas of management: production and distribution, chains of supply and projects, finance and tools of business analysis, relationships with customer and staff [Microsoft Dynamics official website, 2014]. It includes the module «Trade and Logistics» Microsoft Dynamics AX 4.0, which supports the following areas of work departments:

- everyday monitoring: operations for the purchase, sale, other storage operations, work with the specifications;
- setting of storage and nomenclature units;
- organization of the storage infrastructure: managing the placement and packaging of freights for individual storage cells, monitoring of material handling equipment, accounting of tare;
- analysis of inventory flows: obtaining of information about storage movements and stock levels, the use of special reports on logistics, such as ABC-classification, etc .;
- planning of inventory movement: making plans and working with them, calculation of needs the product and using of its results, and forecasting of inventory levels.

Cost of the license of a single workplace is $2200 – 3500.

Advantages:

- the number of concurrent users is from 20 to 5000;
- support of more than 40 languages;

- automates specific and complex business processes (distributed enterprise, holdings, distribution and manufacturing companies, service providers, and so on).

Disadvantages: lack of the transport logistics and cartography; require training to work with the system.

Conclusion: Axapta is actively used in agrobusiness as a software package, which includes a logistics module. However, as regards our LIS, it only partially solves the problem of storage accounting tasks. Due to interfaces it is suitable for management and accounting of the cost of storage of agricultural products, the planning processes of distribution and sales by region.

Oracle Transportation Management

Oracle Transportation Management (developed by the Oracle Corporation, USA) provides a full-featured tool for planning and execution of transportation considering the production capabilities of shippers and third transport companies. This integrated solution allows to combine and to simplify the process of transportation planning, execution, payments to suppliers and customers and to automate business processes within a single application, considering all types of transportation [Oracle official website, 2014].

Oracle Transportation Management (TMS) can be used together with ERP or legacy system Order Management, and any best in its class solution or legacy system Warehouse Management System (WMS). Oracle Transportation Management system is not actually part of the Oracle E-Business Suite, although for some time Oracle has positioned it in such way due to replacing old logistics module by it in OeBS. TMS is also integrated with Oracle Order Management and Oracle Warehouse Management.

It allows solving the following management traffic tasks:

- management of orders for transportation;
- fleet management;
- transportation planning;
- reservation of transportation in carriers, tenders;
- tracking of delivery events;
- automation of solutions about exceptions;
- container optimization;
- minimize the transportation cost.

New versions, starting from Oracle Transportation Management 6.3, include advanced capabilities of fleet management, procurement of transportation services; business analysis in the area of freight transportation; transport planning; management of rail transportation; automation of processes and event management; management and control of freight payment, billing for services and resolution of claims; management of transport and customs documents.

Using two specialized mobile applications TMS also supports key business processes, such as interaction with service providers of organization of transportations, tracking execution of orders/transportations, event management, data access, business analysis of freight transportations on mobile devices.

Cost of the license is $80,000.

Cost of the software is $16,000.

Advantages: improving the quality of transportations: the transparency of terms, reducing the risk of loss of the freight; system transparency for top-management.

Disadvantages: cost – it is impossible to purchase module separately, but only with the package OeBS; insufficient protection: vulnerability of the system allows a remote user to manipulate certain information, to reveal confidential data and to cause denial of service.

Conclusion: product allows managing of transport and transportation planning in agricultural enterprises, in addition to other modules OeBS connects management of storage operations. But it does not take into account the peculiarities of the transport of agricultural freights (harvest season, short harvesting terms and others, which are specified in the previous section).

«1C – Rarus», «Transport logistics and forwarding»

«1C – Rarus», «Transport logistics and forwarding» is a solution for automation of the process of transportation for companies that provide services for the delivery and freight forwarding by various modes of transport: road, rail, air, sea [1C-Rarus official website, 2014]. Works on a platform of «1C – Enterprise». Services are provided for companies with their own transport and for companies that do not have their own vehicles.

Main functional features:

1. Order Management of clients:

- accounting of clients;
- selection of the necessary vehicles and carriers;
- creating the orders for carriers;
- the ability to create templates of standard operations and automatic calculation of the cost of services tariff;
- registration of orders for freight transportation, tracking of future status of freight and its history of changes;
- distribution of freights in containers and container traffic accounting;
- accounting of nomenclature of transported freights;
- accounting of additional characteristics of freights;

- accounting of transactions for freights and orders to the carrier;
- registration of runs of carriers, distribution of freights and containers at runs;
- planning of moving of freights along the route and operational data input about the location of the freights;
- location and shipment of freights from the storages;
- registration of the transfer of freights to the consignee.

2. Fleet management:

- accounting of own and attracted vehicles;
- direction of vehicles in runs;
- control of disposition.

3. Management of mutual settlements:

- accounting of price-lists and tariffs;
- calculation of the cost of transportation;
- creating bills and acts for transport service.

Technical support is carried out in Ukraine.

Cost of the license is $500.

Disadvantages: product does not include a license for the platform «1C: Enterprise 8» - it must be purchased separately; also there is no mapping module, which makes impossible to automate routes.

Conclusion: it covers such areas of logistics, as storage, partly transport and distribution, taking into account the platform «1C: Enterprise 8». It is applied in agribusiness. However, the system does not allow the features of transportation of agricultural freights for solving our tasks.

«Top Logistic»

«Top logistic» is a product of the «TopPlan», developed in Russia. System is designed for automation of planning process of freights delivery by autotransport. The company specializes in the development of electronic maps, databases, and solution of transport tasks [TopPlan official website, 2014].

The main purpose of this product is creation of the most efficient delivery route, considering parameters such as: time interval of delivery, points of delivery address, etc. The program is designed for small, medium and large enterprises. An important advantage of the program is GPS/GLONASS module, which allows controlling traffic in real time and movement the recording of way to archive. This system allows to perform one of the most important task of organization: a comparison of actually performed works with the plan works. The software also has a mapping module and detailed maps of regions and cities (Ukraine, Russia and CIS countries).

The system provides:

- simple and fast integration with the accounting system of the enterprise;
- automation of works connected with distribution of orders between vehicles;
- automatic calculation of routes of orders delivery;
- visualization of addresses and delivery routes on the electronic map;
- creation of optimal order of detour delivery points with the ability to change it.

The system generates:

- database of warehouses for orders shipment;
- database of vehicles with the characteristics of each vehicle;
- database of consignees with addresses that attached to the map;
- database of customer orders with the characteristics of the orders.

The system calculates:

- planned gasoline consumption, kilometrage and working time of each vehicle;
- fuel costs for each run;
- the fuel component of the unit costs of delivery of freight;
- loading factors of vehicles and runs;
- needs for vehicles for delivery of freight.

Cost of license: negotiated price is from $ 1,000 for basic set of functions.

Disadvantages: absence of the module of storage logistics.

Conclusion: It only partially solves the tasks of transport logistics. Logistics industry requires the integration with other software packages.

«INGIT. Business Map»

«INGIT. Business Map», is a product of «INGIT», Russia, which is designed for optimization of freight transport flows, adding the clients on the map and data processing, for solution of analytical and commercial tasks [INGIT official website, 2014].

Business Map can be integrated with any databases (including «1C - Enterprise» for storage operations).

Functional capabilities:

- Creation and supply of client database that is automatically applied to map by the addresses, automatically sorted by the administrative-territorial division and service areas. Databases ACCESS, EXCEL, FOXPRO, PARADOX, LOTUS, DBASE can be used as data sources, which processes Business Map.

- Usage of detailed map allows performing not only logical selections of the client database, but also any space selection (in service areas, near points, etc).
- By client location at the map it is possible to do selection for detour routes for controlling the length, number of clients, and the total load.
- The system of automatic routing of vehicles provides real-route for vehicle for customer service, considering traffic organization, which are selected from the database using any queries or selected on the map, considering the weight and quantitative restrictions.
- Automation of calculations for the freights delivery provides instant calculation of routes for whole vehicle fleet to perform the whole volume of daily orders. Load capacity and transport capacity, requirements of urgency of orders and execution time, optimization of movement including road signs, restrictions on time or length of the routes, for example, to deliver hot bread in points of sale etc. are accounted in the calculations. The system of the results presentation allows to give each driver a printout of his day job, i.e. when, where and what to load, sequence of deliver, and so on. It provides a dynamic view of transport operations for delivery of orders, i.e. at any time it can be seen on the map where the vehicles are located and what they do, it means unloaded, loaded, or follow into the next point.
- Documentation system provides a printout of all forms of documents (lists, bills of lading, waybills, etc.) for customers, selected from the database for any queries or detour routes.
- Multifunctional Information center: it performs quick search of streets, homes, companies, organizations and firms with map display. The open database allows inserting and updating of information by types of activities with simultaneous reference to the map. Information can be supplemented with references to any files (documents, images, etc.), which will be used by appropriate programs.

Cost of license: $600 per a basic set including a server part and a permanent client access license. Additionally it is necessary to buy access key to the cloud GIS server each month.

Disadvantages: inconvenient interface; problems with self-installing of the program and with the transfer to another computer; nonoptimal work of the map.

Conclusion: although it is used in agriculture, it is inappropriate to use this system for solving of the required tasks. It covers them partly – only in the part of the transport and distribution logistics.

4. Comparative characteristic of software products, depending on the tasks that must be solved by LIS

After more detailed description of software for logistics was performed, it is necessary to know the functional suitability of each program and to do comparative analysis (Table 1).

Table 1. Summary table of comparative analysis of functional suitability of software packages

Program product / Tasks	Optimum GIS	Chronoma p	ArcGIS	Axapta	OTM	1C – Rarus	Top Logistic	INGIT. Business
planning of optimal delivery routes	yes	partly	yes	no	yes	no	yes	yes
creating of operational planning schedules of transport movement	partly	partly	partly	no	partly	no	partly	partly
dynamic of planning of transport usage	partly	partly	partly	no	partly	no	partly	partly
accounting of set of conditions in planning of routes	partly	partly	partly	no	partly	no	partly	partly
forecasting, distribution and planning of material resources	no	no	no	yes	yes	partly	no	partly
accounting of restrictions of delivery and features of freights	partly	no	partly	no	partly	no	no	partly
planning and management of maintenance and repair of rolling stock	no	no	no	yes	yes	no	yes	no
planning in stationary conditions and in conditions with infrastructure changes	partly	partly	partly	partly	partly	no	partly	partly
inventory management in storages of vehicles and organization of storing materials at storages	no	no	no	yes	yes	yes	partly	partly
accounting of deadlines storage of resources	no	no	no	partly	partly	no	no	no
planning the usage of additional resources	partly	no	no	no	yes	yes	partly	partly
developing of programs for economy of material resources, usage of transport and control over their implementation	partly	partly	no	partly	yes	yes	partly	partly

5. Conclusions

All software solutions provide wide coverage of management functions, including logistics and inventory management. But neither product provides complex solution of our tasks for agricultural. This is associated with the fact that these software systems are more oriented to large corporate structure, and not every medium-sized company can afford to purchase them. The other reason is the situation with the intricacies of their introduction, the payback period, the availability of experts for teaching about system usage features, features of the enterprise. Domain is not completely investigated and requires new algorithms for solving such tasks as:

- route planning considering randomness in time, the type of vehicle;
- route planning considering real road network, weather conditions, etc., and pave roads that are not present on the map;
- route planning considering delivery restrictions and features of agricultural freights;
- planning in stationary conditions and in the conditions of dynamically changing states of transport fields, roads, etc.;
- cost reduction by decreasing fuel consumption, kilometrage, transportation costs, travel time, the number of involved vehicles, the waiting time.

For solving such types of tasks it is necessary to develop new information technologies and program solutions.

Bibliography

[1C-Rarus official website, 2014] 1С-Рарус : Транспортная логистика и экспедирование [Online]. – Available: http://rarus.com.ua/transp/1c-rarus-translogistic-ukraina/

[ArcGIS official website, 2014] ArcGIS. [Online]. – Available: https://www.arcgis.com/features/

[CDC official website, 2014] Автоматизация транспортной логистики, GPS/ГЛОНАСС мониторинг. ОПТИМУМ ГИС. [Online]. – Available: http://www.cdc.ru/solutions/optimum-gis/

[INGIT official website, 2014] INGIT. Business Map. [Online]. – Available: http://www.ingit.ru/businessmap/

[Lytvynov, 2014] Lytvynov V.V., Mykhailiuk I.V., Posadska A.S. Functional features of dispatching control centre for automatic control system of agricultural enterprise. Науковий журнал. Математичні машини і системи, Київ, 2014, №.3, 67-77.

[MapInfo official website, 2014] ChronoMap. [Online]. – Available: http://www.mapinfo.ru/product/chronomap

[Microsoft Dynamics official website, 2014] Microsoft Dynamics ERP. [Online]. – Available: http://www.microsoft.com/ru-ru/dynamics/erp.aspx

[Oracle official website, 2014] Oracle Transportation Management. [Online]. – Available: http://www.oracle.com/us/products/applications/ebusiness/logistics/018756.htm

[TopPlan official website, 2014] TopLogistic. [Online]. – Available: http://www.toplogistic.ru/

[Бальченко, 2013] Бальченко И.В., Литвинов В.В., Клименко В.П. Особенности построения автоматизированной системы управления сельскохозяйственным предприятием. Науковий журнал. Математичні машини і системи, Київ, 2013, № 4, 82-94.

[Бочкарев, 2013] Бочкарев А.А. Методика выбора информационной системы для автоматизации планирования автотранспортной доставкой грузов. Проблемы современной экономики, 2013, № 1 (45), 117-119.

[Гончаров, 2012] Гончаров П.П., Едаков А.Е. Развитие логистики на предприятиях АПК. Известия оренбургского государственного аграрного университета, 2012, № 1-1 / том 33.

[Костышева, 2013] Костышева Я.В. Эффективность применения программных обеспечений в области транспортной логистики. Экономикс, № 1, 2013, 47-54.

[Разгуляев, 2010] Разгуляев В. Автоматизация планирования и учёта перевозок предприятия. Материалы научно-практической конференции «Логистика и управление цепями поставок: эффективное взаимодействие участников (субъектов) транспортного процесса». 2010.

[Сергеев, 2008] Сергеев В.И., Григорьев М.Н., Уваров С.А. Логистика: информационные системы и технологии. Учебно-практическое пособие. Москва. Альфа-Пресс, 2008.

Authors' Information

Vitalii Lytvynov *– Dr. Sc., Prof. Chernihiv, National University of Technology, 95, Shevchenko street, Chernihiv-27, Ukraine, 14027; vlitvin@ukrsoft.ua*

Major Fields of Scientific Research: modeling of complicated systems, computer-aided management systems, decision support systems

Alina Posadska *– Ph.D. Student, Chernihiv National University of Technology, 95, Shevchenko street, Chernihiv-27, Ukraine, 14027; alinka.posadskaya@gmail.com*

Major Fields of Scientific Research: technologies and methods of solving of logistic tasks, software solutions of network planning

PASSIVE STEGANALYSIS OF MULTIDOMAIN EMBEDDING METHODS

Dmytro Progonov, Serhii Kushch

***Abstract**: The paper is devoted to analysis of effectiveness the usage of modern statistical model of digital image for revealing the stego images with data, embedded in transformation domain of cover images. It is considered the case of applying of standard Subtractive Pixel Adjacency Matrix model for detection the presence of stegodata, hidden with usage of various type the cover image transformation – Discrete Cosine Transform, Discrete Wavelets Transform, Singular Value Decomposition. It is established that accuracy of stego image detection by usage of the model rises by increasing the amount of stage the cover image processing and significantly depends on type of embedded stegodata.*

***Keywords**: passive steganalysis, multistage embedding methods, digital images.*

***ACM Classification Keywords**: D.4 Operating Systems – Security and protection – Information flow controls.*

Introduction

During the last days of computer systems widespread usage of high speed communication systems (CS), integration of government agencies and private corporations computers systems into global network, usage of various types the communication services led to corresponding revision of methodology the attacks of malefactors and terrorists [Tallinn, 2013]. Significant part of success these attacks depends on the reliable communication between the intruders for data transmission and coordination of actions. In most cases such communication channels are embedded in existed information flows in various types the communication services, such as social networks, multimedia files sharing, Voice-over-IP services, with usage of steganographic systems (SS) [The Cisco, 2014]. Therefore, the early detection and counteraction the hidden messages (stegodata) transferring in communication services are important task today.

As cover media for stegodata are widely used difference types of multimedia files, in particular digital images (DIs), which is explained by high redundancy of theirs representation in digital form [Fridrich, 2010]. Existed methods of image steganography can be divided in two groups – embedding in spatial domain (LSB-methods) and in the transformation domain (TD) [Katzenbeisser, 2000]. LSB methods allow minimizing the distortion of cover image parameters by stego images forming, but has relatively low robustness to any alteration of stego image by DIs transferring in CS. Message hiding in TD is based on usage the different types transformations of cover images by stegodata embedding. It allows considerably increasing the robustness of obtained stego images to active steganalysis by preservation the fixed distortion of cover image’s parameters.

For revealing the stego images with data, embedded in spatial and transformation domains of cover images, were proposed effective approaches, based on statistical [Pevny, 2008; Fridrich, 2012], multifractal [Progonov, 2014a, 2014b, 2014d], variogram [Progonov 2014c] and spectral [Doroshenko, 2014] analysis. In spite of “universality” the statistical methods, in most cases they are used for revealing the stego images with data, embedded only in spatial domain or JPEG domain of cover images. Therefore it is represented the interest to investigate the effectiveness of applying statistical steganalysis for discerning the stego images with data, hidden in TD.

In the paper we investigated the effectiveness of usage the statistical model of images, based on Markov features the DIs, for detection the presence of stegodata, hidden in various TD. Obtained results can be used by creation of universal stegodetectors for revealing the stego image with data, embedded in spatial as well as difference transformation domains of digital images.

Related Works

For revealing the stego images with data, embedded in spatial domain of DIs, there were proposed structural analysis method [Dumitrescu, 2002] and approaches, based on applying the statistical models (SMs) of DI [Pevny, 2010b; Kodovsky, 2009]. In most cases SM are created for realization the targeting attack on specified SS, which limits further usage of proposed SM for revealing the another methods of stegodata embedding. For overcome mentioned limitation it was proposed [Fridrich, 2012] to merge the separate SMs into the rich models (RMs) of DI. It allows successfully attack the modern highly undetectable embedding algorithms (for instance, HUGO algorithm [Pevny, 2010a]), which has been impossible with usage of simple SMs. Limitation of practical usage the RMs is high dimensionality of features space (for example, 34671 features for SRM model [Fridrich, 2012]), which leads to complication of stegodetector tuning procedure. Due to this, further improvements of RM are carried out by optimization the used feature space or development the alternative statistical models of DIs [Holub, 2013b, 2015].

Alternative approach to create the highly undetectable embedding algorithms (HUEA) is message hiding in TD of cover image, in particular in spectral domain of DI (for instance, WOW algorithm [Holub, 2012], UNIWARD method [Holub, 2013a]). Rich models, proposed for revealing of such HUEA, are based on assumption the stegodata hiding in JPEG-domain of DIs [Kodovsky, 2009, 2012b], which led to narrowing of application domain for these models. Therefore it is represented the interest to investigate the effectiveness of applying the standard spatial domain-based RMs for revealing the stegodata, embedding with usage both spectral (Discrete Cosine Transform (DCT), Two-Dimensional Discrete

Wavelets Transform (2D-DWT)) and special (Singular Value Decomposition, SVD) transformation of the cover images.

The Goal and Contribution

The goal of paper is investigation of effectiveness the application of Subtractive Pixels Adjacency Matrix (SPAM) model for revealing the stego image with data, embedded with usage of one-stage and multi-stage methods in the spectral and singular value domains of the cover images.

Multidomain Data Embedding Methods

Historically, the first methods, proposed for message embedding in spatial domain, were based on substitution the least significant bits of pixels by stegodata [Fridrich, 2010]. These methods allows embedding the message with volume up to 1/8 of cover image's size, but result in specific distortion of image parameters – changing the statistics of bit value distribution (chi-square attack [Westfeld, 1999]) or local correlation among neighboring pixels (Sample Pairs Analysis [Dumitrescu, 2002], RS Analysis [Fridrich, 2004]). Modern approaches to message hiding with applying of LSB-methods based on adaptive selection of image context for stegodata embedding and usage the encoding systems for minimization the amount of changed pixels [Fridrich, 2010].

For increase the robustness of the stego images to possible alteration or intentional changes by image transmission in CS, Zhao and Koch proposed to use the peculiarities of DCT [Zhao, 1995]. Based on this work, the separate class of methods for stegodata embedding in DIs with usage of classical spectral as well as special transformations was developed.

Choice the type of cover image transformation depends on requirements to SS – usage of spectral transformation (for instance, Discrete Wavelets or Cosine Transforms) allows increase the endurance of formed stego images to applying the standard transformation, such as lossy compressions, while special types of transformation (for example, Singular Value Decomposition) gives opportunity to decrease the distortion of cover image parameters.

In the work we analyzed the methods, which are based on usage both groups of cover image transformations, as well as the case of one-stage and multistage message embedding, when several transformation are used simultaneously.

The list and parameters of investigated embedding methods are represented in table 1. Message embedding was provided by weighted summation of transformation coefficient the cover image K_{cover} and stegodata K_{data} in specified transformation domain:

$$K_{stego} = K_{cover} + G \times K_{data}, \qquad (1)$$

where G – weighted coefficient, which is used for variation the energy of stegodata. Values of the coefficient G were changed from $G_{\min}$ (lower bound of stegodata reconstruction on receiver’s side the SS) to $G_{\max}$ (appearance the visual distortion of cover image by message hiding) with step Δ_G (Table 1). Cover image payload is determined as fraction of changed coefficients by message hiding to total amount of the transformation coefficients. By stego images forming each color channels of cover image and stegodata were processed independently.

Table 1. Analyzed methods the message embedding in transformation domain of digital images

Authors	Cover image processing			Stegodata processing	Weighted coefficient G		
	1st stage	2nd stage	3rd stage		$G_{\min}$	$G_{\max}$	Δ_G
Dey (2011)	2D-DWT	–	–	2D-DWT	0.02	0.08	0.02
Agarwal (2006)	SVD	–	–	SVD	0.02	0.08	0.02
Joseph (2013)	2D-DWT	SVD	–	SVD	0.1	2	0.5
Khan (2013)	2D-DWT	DCT	SVD	SVD	0.5	4	1

Approximation W_φ and detailed W_ψ coefficients of 2D-DWT transform of grayscale image $I_{x,y}$ with size $M \times N$ (pixels) were calculated according to further formulae [Gonzalez, 2008]:

$$W_\varphi(j,m,n) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I_{x,y} \times \varphi_{j_0,m,n}(x,y),$$

$$W_\psi^i(j,m,n) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I_{x,y} \times \psi_{j,m,n}^i(x,y), i \in \{H,V,D\},$$

where

$$\varphi_{j,m,n}(x,y) = 2^{j/2} \times \varphi(2^j x - m) \otimes \varphi(2^j y - n),\ \psi_{j,m,n}^H(x,y) = 2^{j/2} \times \psi(2^j x - m) \otimes \varphi(2^j y - n),$$

$$\psi_{j,m,n}^V(x,y) = 2^{j/2} \times \varphi(2^j x - m) \otimes \psi(2^j y - n),\ \psi_{j,m,n}^D(x,y) = 2^{j/2} \times \psi(2^j x - m) \otimes \psi(2^j y - n),$$

correspondingly, two-dimensional scaling function $\varphi_{j,m,n}(x,y)$ and wavelets $\psi_{j,m,n}^i(x,y)$; $\varphi(x), \psi(x)$ – one-dimensional scaling and wavelet functions; $\otimes$ – Cartesian product; j_0, j – initial and current decomposition levels; m,n – spatial shift parameters for two-dimensional scaling and wavelet functions.

Inverse DWT was calculated according to formula [Gonzalez, 2008]:

$$I(x,y)=\frac{1}{\sqrt{MN}}\times\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left[W_{\varphi}(j_0,m,n)\times\varphi_{j_0,m,n}(x,y)+\sum_{i\in\{H,V,D\}}\sum_{j=j_0}^{+\infty}W_{\psi}^{i}(j,m,n)\right].$$

According to Dey, Joseph and Khan methods the detailed coefficient were used for message hiding in cover image, which is explained by features of human vision – relatively low sensitivity to slight changes of fine details the images. As basic functions of 2D-DWT were used the Haar wavelet and corresponding scaling function.

Direct and inverse DCT of grayscale image $I_{x,y}$ with size $M\times N$ (pixels) were calculated according to further formula [Oppengeim, 2010]:

$$T(u,v)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}I_{x,y}\times r(x,y,u,v);I_{x,y}=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}T(u,v)\times s(x,y,u,v),$$

$$r(x,y,u,v)=s(x,y,u,v)=\beta(u)\times\beta(v)\times\cos\left[\frac{(2x+1)u\pi}{2M}\right]\times\cos\left[\frac{(2y+1)v\pi}{2N}\right],$$

$$\beta(u)=\begin{cases}\sqrt{1/M},u=0;\\ \sqrt{2/M},u=1,2\ldots(M-1);\end{cases}\quad \beta(v)=\begin{cases}\sqrt{1/N},\mathrm{v}=0;\\ \sqrt{2/N},\mathrm{v}=1,2\ldots(N-1);\end{cases}$$

where $r(x,y,u,v),s(x,y,u,v)$ – correspondingly, kernels of direct and inverse DCT, $\beta(u),\beta(v)$ – normalization multipliers.

SVD of grayscale image $I_{x,y}$ with size $M\times N$ (pixels) were provided according to formula [Murphy, 2012]:

$$I_{x,y}=U_{M\times M}\times S_{M\times N}\times V_{N\times N}^{T},$$

where U,V – orthonormal matrix of left and right eigenvectors; S – diagonal matrix, which contains the singular values of matrix $I_{x,y}\times I_{x,y}^{T}$. Stegodata embedding was provided with usage of eigenvalues due to ambiguous the eigenvector reconstruction (up to theirs permutation) on the receiver’s side of SS.

SPAM Model of Digital Images

Consequence of embedding stegodata in cover images with usage of any steganographic methods is alteration of cover parameters. Due to this, significant part of modern steganalysis methods is based on analysis the changing of cover image parameters and further creation the cluster of image’s characteristics, which changes at most by stego image forming. It should be mentioned that in most cases providing of targeting steganalysis is impossible due to absence or scantiness the priory information about the embedding domain or hiding algorithm. Therefore modern approach for stego image detection is usage of cover RMs (CRMs) – consolidation of several simple SM of cover images, which is based on peculiarities of DI (for instance, Markov features) or specific of image parameters alteration by message hiding.

Existed CRMs can be divided into two groups – spatial-domain and JPEG-domain based models. These models were developed for targeting steganalysis of modern embedding methods (for instance, CC-PEV model for attack the YASS algorithm [Kodovsky, 2009]) or creation the universal (blinf) stegodetectors for revealing the stego images in case of absence the a priory information about the embedding method (for example, [Kodovsky, 2012b]). Limitation of known JPEG-domain based models is theirs ability for detection only the specific types of hiding algorithm, while spatial-domain based models (SDBMs) also give opportunity to provide the blind steganalysis. Therefore it is represented the interest to use the SDBMs for revealing the steganograms with data, embedding in various TD.

In the work we investigated the efficiency of well-known Subtractive Pixel Adjacency Matrix (SPAM) model [Pevny, 2010b]. SPAM model is based on usage the first and second order Markov chains (MCs) for modeling the dependencies between difference $D_{x,y}$ of adjacency pixels in grayscale image $I_{x,y}$. For instance, difference between brightness of horizontally adjacent pixels can be represented as:

$$D^{\rightarrow}_{x,y} = I_{x,y} - I_{x+1,y}.$$

Then parameters $M^{\rightarrow}_{u,v}$ ($M^{\rightarrow}_{u,v,w}$) of first (second) order of MC were calculated according to further formulae [Pevny, 2010b]:

$$M^{\rightarrow}_{u,v} = \Pr\left(D^{\rightarrow}_{x+1,y} = u \mid D^{\rightarrow}_{x,y} = v\right), M^{\rightarrow}_{u,v,w} = \Pr\left(D^{\rightarrow}_{x+2,y} = u \mid D^{\rightarrow}_{x+1,y} = v \mid D^{\rightarrow}_{x,y} = w\right), u,v,w \in \{-T,...T\},$$

where $\Pr(A)$ – probability of event A; T – specified threshold, which is used for limit the variability of $D_{x,y}$ values. Consolidate the parameters of MC, we can write the whole SPAM model as:

$$\begin{cases} F_{1,2...k} = \frac{1}{4}\times\left[M^{\rightarrow}_{*} + M^{\leftarrow}_{*} + M^{\uparrow}_{*} + M^{\downarrow}_{*}\right]; \\ F_{k+1,k+2...2k} = \frac{1}{4}\times\left[M^{\searrow}_{*} + M^{\nwarrow}_{*} + M^{\swarrow}_{*} + M^{\nearrow}_{*}\right]. \end{cases}$$

where $k = (2T+1)^2$ and $k = (2T+1)^3$ for first and second order MC correspondingly.

For increasing the performance of SPAM model, according to recommendation [Pevny, 2010b], in the work were used the second-order MC for modelling the difference between adjacent pixels with threshold $T = 3$. Therefor the dimensionality of features space for grayscale cover image was equal to $2\times(2T+1)^3 = 2\times 7^3 = 686$.

Results

Analysis of effectiveness the SPAM model for revealing the stego images with data, embedded in TD, were provided with usage of cover images (JPEG, True Color) packet MIRFlickr-25k [Huiskes, 2008]. For training and testing of stegodetector were used the subset of 2,500 pseudo randomly selected and scaled DI from packet. As stegodata were used three DI – engine’s draft, map and portrait. Characteristics of the stegodata are represented in Table 2:

Table 2. Characteristics of used test digital images and stegodata

Characteristics	Engine’s draft	Map	Portrait
Resolution, pixels	567×463	800×800	565×850
Color system	RGB		
Format	BMP		

By investigation, cover image payloads were changed from 5% to 25% with step 5% and from 25% to 95% with step 10%. Weighted coefficients G, for each investigated embedding method, were changed from $G_{\min}$ up to $G_{\max}$ with step Δ_G (Table 1).

The stegodetector was training with usage of half the test packet. Testing the tuned stegodetector was provided on the remained half of test packet. Recognition of stego images by stegodetector was provided with usage of ensemble classifier [Kodovsky, 2012a]. As base classifier was used the Fisher’s Linear Discriminant (FLD), which was tuned for minimization of total detection error P_E on training subset the test packet:

$$P_E = \min_{P_{FA}} \frac{1}{2}\left[P_{FA} + P_{MD}\left(P_{FA}\right)\right],$$

where P_{FA}, P_{MD} denote the probabilities of false alarm and missed detection respectively. Assessment of P_{FA} and P_{MD} was provided according to bootstrap estimation algorithm [Kodovsky, 2012a] by training each base classifier B_l on pseudo random selected subset of training set

$$\mathrm{X}_l = \left\{ x_m^{(D_l)}, \bar{x}_m^{(D_l)} \right\}_{m \in \mathfrak{M}_l^b},$$

where $x_m, \overline{x}_m$ – training samples; $D_l, l \in \{1,2...d\}$ – pseudo randomly selected subset of features from general feature space with dimensionality d ; $\mathfrak{M}_l^b$ – bootstrap sample of $\{1,2...N^{trn}\}$; N^{trn} – amount of test cover images at training stage.

The total detection error P_E (out-of-bag (OOB) error) for stegodetector after training phase was computed according to formula:

$$P_{E-OOB}^{(n)} = \frac{1}{2N^{trn}} \sum_{m=1}^{N^{trn}} \left[B^{(n)}(x_m) + 1 - B^{(n)}(\overline{x}_m) \right].$$

Analysis of accuracy the stego image detection was provided with usage of SPAM model was provided for two cases – with usage of all or separate stegodata for stegodetector training and testing. Investigation was provided for grayscale (separate color channels of test DIs) and true color images. Estimation of mean value and variance of the OOB-error P_E was provided by repeating the training and testing stage 10 times.

Results of testing the performance of stegodetector, tuned with usage of SPAM model, by message hiding in TD are represented at Figure 1 and Table 3.

It should be mentioned the unexpected results for the Dey, Joseph and Khan methods – total detection error P_E depends only on weighted coefficient G, in other words on stegodata energy (Figures 1b - 1d). It can be explained by ability of 2D-DWT to separate the details on DI depends on theirs orientation. Due to this it is possible to minimize the distortion of Markov features by selection of detailed coefficient (direction of fine details) for message hiding.

Secondly, usage of SVD (Figure 1a) allows considerably decrease the precision of stego image detection in comparison with 2D-DWT cases (Figures 1b - 1d). Also, the high level of OOB for Agarwal method preserved for wide range of cover image payloads and weighted parameter G values, which indicates about the relatively low effectiveness of usage the SPAM model it this case. Obtained results are explained with usage statistical interpretation of SVD transform – opportunity to decompose the DI on components with maximum variance [Murphy, 2012].

Usage of cover image’s components with higher variance for message embedding allows decrease the changes of difference the brightness of adjacent pixels and, correspondingly, alteration the parameters of MC and SPAM model.

Results, represented at Figure 1, allow us estimating the minimum OOB error for SPAM model due to usage of all color channels and types of stegodata at training phase of stegodetector tuning. So, let see the results for more “realistic” situation, when only part of available information can be used for stegodetector adjustment (Table 3):

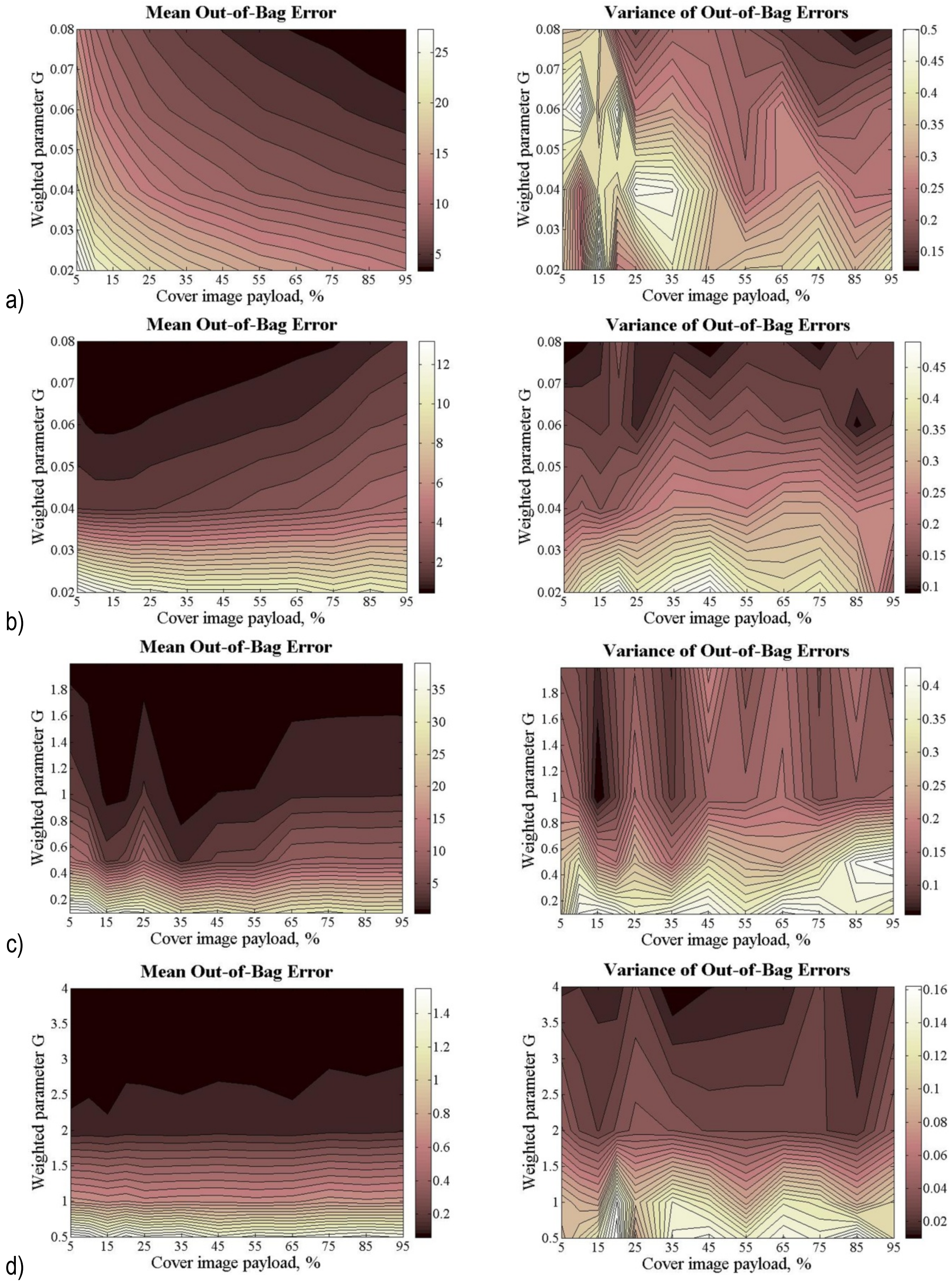


Figure 1. Mean value and variation of total detection error P_E for stegodetector by usage of SPAM model and all cover image color channels and stegodata type. Message was embedded according to: (a) – Agarwal method; (b) – Dey method; (c) – Joseph method; (d) – Khan method

Table 3. Mean values and variance of total detection error P_E for investigated methods (Maximum OOB-error / Variance of OOB-error / Minimum OOB-error)

	All color channel	Separate color channel		
		Red color	Green color	Blue color
Agarwal method				
All stegodata	28.78 / 0.52 / 3.37	31.97 / 0.53 / 5.51	33.45 / 0.54 / 4.93	35.07 / 0.44 / 5.36
Stegodata type "Draft"	30.67 / 0.81 / 3.17	32.86 / 1.07 / 4.36	34.06 / 0.96 / 3.82	36.53 / 1.07 / 4.62
Stegodata type "Map"	30.77 / 1.06 / 3.28	33.08 / 0.94 / 5.13	34.64 / 0.85 / 4.16	36.73 / 0.997 / 4.85
Stegodata type "Portrait"	30.46 / 0.87 / 4.52	32.88 / 0.90 / 7.35	34.66 / 0.86 / 6.14	35.99 / 0.92 / 7.03
Dey method				
All stegodata	13.70 / 0.51 / 0.50	26.90 / 0.96 / 1.70	20.80 / 0.71 / 1.20	27.00 / 0.85 / 1.30
Stegodata type "Draft"	17.00 / 0.88 / 1.20	26.40 / 0.76 / 2.00	19.70 / 0.78 / 1.40	26.00 / 0.72 / 1.70
Stegodata type "Map"	17.70 / 0.88 / 0.80	26.80 / 0.80 / 1.70	20.90 / 0.87 / 1.30	27.10 / 1.03 / 1.40
Stegodata type "Portrait"	18.20 / 1.02 / 1.30	27.60 / 1.01 / 2.10	20.90 / 0.71 / 1.50	27.30 / 1.05 / 1.80
Joseph method				
All stegodata	40.67 / 0.43 / 0.49	43.19 / 0.40 / 0.64	43.14 / 0.61 / 0.82	44.09 / 0.45 / 0.63
Stegodata type "Draft"	38.86 / 1.08 / 0.37	40.64 / 0.78 / 0.22	40.00 / 0.91 / 0.40	41.80 / 0.81 / 0.31
Stegodata type "Map"	46.28 / 0.83 / 0.26	47.92 / 0.66 / 0.28	47.23 / 0.90 / 0.23	46.94 / 0.96 / 0.34
Stegodata type "Portrait"	46.84 / 0.87 / 0.83	47.50 / 0.62 / 1.16	47.15 / 0.75 / 1.09	47.27 / 0.81 / 0.59

	All color channel	Separate color channel		
		Red color	Green color	Blue color
Khan method				
All stegodata	1.61 / 0.17 / 0.07	2.10 / 0.19 / 0.08	1.83 / 0.13 / 0.06	1.94 / 0.12 / 0.07
Stegodata type "Draft"	1.67 / 0.18 / 0.00	2.17 / 0.27 / 0.00	1.78 / 0.25 / 0.00	1.96 / 0.21 / 0.00
Stegodata type "Map"	1.83 / 0.24 / 0.00	2.58 / 0.24 / 0.01	2.26 / 0.26 / 0.00	1.99 / 0.17 / 0.01
Stegodata type "Portrait"	2.21 / 0.29 / 0.11	2.73 / 0.34 / 0.08	2.69 / 0.25 / 0.10	2.77 / 0.38 / 0.11

First of all, it should be mentioned that values of OOB errors are significantly variance for various color channels (Table 3) – OOB error is minimal for green channel and maximum for blue channel (Agarwal, Dey and Joseph methods) or red channel (Khan method). Obtained results are explained by procedure of DI acquisition in digital camera or scanners – presence of demosaicing (debayering) stage, when the equalization of energy the DI spectral components according to peculiarities the human vision is carried out [Fridrich, 2010]. It is realized by averaging the values for adjacent green subpixels, which leads to corresponding suppression of noise. Message hiding leads to distortion of statistics parameters for green color channel, which is registered by corresponding change of results for SPAM model.

Also, the values of OOB errors for stegodata type "Portrait" are higher for all considered methods in comparison with other types of stegodata (Table 3). Such "imbalance" in obtained results is explained by substantially lesser amount of fine details for Portrait-stegodata, which leads to corresponding decreasing the number of changed pixels by stego image forming.

Conclusion

On the basis on conducted analysis of OOB errors by usage the SPAM model for detection the stego images with data, embedded with applying of various transforms, it is established that:

1. Effectiveness of SPAM model significantly depends on amount of stage the cover image processing by stegodata hiding – the lowest OOB-errors are achieved in case of usage the multistage embedding methods. Increase the OOB error took place by applying of SVD for message hiding (Agarwal and Joseph embedding methods), which explained by usage the image components with the higher variance for stegodata hiding;

2. Range of OOB errors value in case of 2D-DWT usage for stego images forming depends only on weighted parameter G. Due to this it is impossible to provide the quantitative steganalysis with usage of tuned stegodetector – estimation of cover image payload by analysis of variation of changes the parameters of SPAM model;

3. Considerable influence on level the OOB error has type of used stegodata and color channel of cover image by message hiding. Therefore it is recommended to provide the adjustment of stegodetectors with usage as much as possible testing message and usage both grayscale as well as true color images for increase the accuracy of steganogram discerning.

Acknowledgements

The paper is published with partial support by the project ITHEA XXI of the ITHEA ISS (www.ithea.org) and the ADUIS (www.aduis.com.ua).

Bibliography

[Agarwal, 2008] Agarwal R., Santhanam M.S. Digital watermarking in the singular vector domain. International Journal of Image and Graphics, 2008. Vol. 8. pp. 351-362.

[Dey, 2011] Dey N., Roy A.B., Dey S. A novel approach of color image hiding using RGB color planes and DWT. International journal of computer application, 2011. Vol. 36, No.5. pp.19-24.

[Doroshenko, 2014] Doroshenko A., Kushch S., Progonov D. Spectral characteristics of steganograms. Proceedings of VII International conference "Suchasni problemy i dosyagnennya v galuzi radiotechniky, telekomunikatiy ta informaciynuh tehnologiy". pp. 327-328 (in Ukrainian).

[Dumitrescu, 2002] Dumitrescu S., Wu X., Memon N.D. On steganalysis of random LSB embedding in continuous-tone images. Proceedings IEEE, International Conference on Image Processing, ICIP 2002. Rochester, NY, USA, September22-25 2002. pp. 324-339.

[Fridrich, 2004] Fridrich J., Goljan M., Du R. Detection LSB steganography in color and grayscale images. IEEE Multimedia, Special Issue on Security, 2001. Volume 8, Issue 4. pp. 22-28.

[Fridrich, 2010] Fridrich J. Steganography in Digital Media: Principles, Algorithms and Applications. Cambridge University Press, New York, USA. 2010. 437 p.

[Fridrich, 2012] Fridrich J., Kodovsky J. Rich models for steganalysis of digital images. IEEE Transactions on Information Forensics and Security, 2012. Vol. 7, Issue 3. pp. 868-882.

[Gonzalez, 2008] Gonzalez R., Woods R. Digital Image Processing. International version 3rd edition. Pearson Education Press, 2008. 1103 p.

[Holub, 2012] Holub V., Fridrich, J. Designing steganographic distortion using directional filters. In Fourth IEEE International Workshop on Information Forensics and Security, December 2–5, 2012.

[Holub, 2013a] Holub V., Fridrich, J. Digital image steganography using universal distortion. Puech, W., Chaumont, M., Dittmann, J., and Campisi, P., eds. In 1st ACM IH&MMSec. Workshop, June 17–19, 2013.

[Holub, 2013b] Holub V., Fridrich J. Random Projections of Residuals for Digital Image Steganalysis. IEEE Transactions on Information Forensics and Security, 2013. Vol. 8, Issue 12. pp. 1996-2006.

[Holub, 2015] Holub V., Fridrich J. Phase-Aware Projection Model for Steganalysis of JPEG Images. Proc. SPIE, Electronic Imaging, Media Watermarking, Security, and Forensics XVII, San Francisco, CA, February 8–12, 2015.

[Huiskes, 2008] Mark J. Huiskes, Michael S. Lew. The MIR Flickr Retrieval Evaluation. Proceedings of the 2008 ACM International Conference on Multimedia Information Retrieval, Vancouver, Canada. ACM Press, New York, NY, USA. DOI 10.1145/1460096.1460104.

[Joseph, 2013] Joseph A., Anusudha K. Robust Watermarking Based on DWT-SVD. International Journal on Signal & Image Security, 2013. Issue 1, Vol. 1.

[Katzenbeisser, 2000] Katzenbeisser S., Petitcolas P. Information Hiding Techniques for Steganography and Digital Watermarking. Artech House, Boston, USA. 2000. 237 p.

[Khan, 2013] Khan M.I., Rahman M., Sarker I.H. Digital Watermarking for image Authentication Based on Combined DCT, DWT and SVD Transformation. International Journal of Computer Science Issues, IJCSI, 2013. Volume 10, Issue 3, No. 1. pp. 223-230.

[Kodovsky, 2009] Kodovsky J., Fridrich J. Calibration revisited. In J. Dittmann,S. Craver, and J. Fridrich editors. Proceedings of the 11th ACM Multimedia and Security Workshop, Princeton, NJ, September 7–8, 2009.

[Kodovský, 2012a] Kodovský J., Fridrich J., Holub V. Ensemble Classifiers for Steganalysis of Digital Media. IEEE Transactions on Information Forensics and Security, 2012. Vol. 7, No. 2. pp. 432-444.

[Kodovsky, 2012b] Kodovsky J., Fridrich J. Steganalysis of JPEG Images Using Rich Models. Electronic Imaging, Media Watermarking, Security, and Forensics. Proceedings of 14th SPIE conference. San Francisco, California, USA. January 23–25, 2012. Vol. 8303. pp. 1-12.

[Murphy, 2012] Murphy Kevin P. Machine Learning: A Probabilistic Perspective. Massachusetts Institute of Technology Press, 2012. 1071 p.

[Oppenheim, 2010] Oppenheim Alan V., Shaffer Ronald W. Discrete-Time Signal Processing. 3rd edition. Pearson Education Press, 2010. 1046 p.

[Pevny, 2010a] Pevný T., Filler T., Bas P. Using high-dimensional image models to perform highly undetectable steganography. In R. Böhme and R. Safavi-Naini, editors. Information Hiding, 12th International Workshop, Calgary, Canada, June 28–30, 2010. Springer-Verlag, New York. Volume 6387 of Lecture Notes in Computer Science, pages 161–177,

[Pevny, 2010b] Pevny T., Bas P., Fridrich J. Steganalysis by Subtractive Pixel Adjacency Matrix. IEEE Trans. on Information Forensics and Security, 2010. Vol. 5, Issue 2. pp. 215-224.

[Progonov 2014c] Progonov D., Kushch S. Variogram Analysis of Steganograms. Proceeding of 3rd International Conference "Zahyst informatii I bezpeka informatiynuh system", 05-06 Juni, 2014, Lviv, Ukraine. pp. 84-85. (in Ukrainian).

[Progonov 2014d] Progonov D., Kushch S. On the Multifractal Analysis of the Steganograms. Proceeding of the 4th International conference "Theoretical and Applied Aspects of Cybernetics", November 24-28, 2014, Kyiv, Ukraine pp. 32-38.

[Progonov, 2014a] Progonov D. O., Kushch S. M. Revealing of steganograms with data, which are hidden in transformation domain of digital images. Visn. NTUU KPI, Ser. Radioteh. radioaparatobuduv., 2014. No. 57, pp. 128-142. (in Ukrainian).

[Progonov, 2014b] Progonov D., Kushch S. Identification the type of transformations, used by message hiding in digital images. Proceeding of 16th International conference "System Analysis and information Technologies", My 26-30, 2014, Kyiv, Ukraine. pp. 435-436. (in Russian).

[Tallinn, 2013] Tallinn Manual On the International Law Applicable to Cyber Warfare. Ed. Michael N. Schmitt. Cambridge University Press, 2013, 302 p.

[The Cisco, 2014] The Cisco 2014 Annual Security Report. San Jose, California, USA. January, 2014.

[Westfeld, 1999] Westfeld A., Pfitzmann A. Attacks on Steganographic Systems Breaking the Steganographic Utilities EzStego, Jsteg, Steganos, and S-Tools—and Some Lessons Learned. Proceedings of Third International Workshop, IH`99. Dresden, Germany, 1999. pp. 61-71.

[Zhao, 1995] Zhao J., Koch E. Embedding Robust Labels into Images for Copyright Protection. Proceedings of the Int. Congress on Intellectual Property Rights for Specialized Information, Knowledge and Techniques. Munich-Vienna, 1995. Verlag. pp.242-251.

Authors' Information

Dmytro Progonov *– the 2nd year postgraduate student, the Assistant, Faculty of Information Security, Institute of Physics and Technology, National Technical University of Ukraine "Kyiv Polytechnic Institute"; Postal Code 03056, Prospect Peremohy, 37, Kyiv, Ukraine; e-mail: progonov@gmail.com.*

Major Fields of Scientific Research: Digital Media Steganography, Advanced Signal Processing, Machine Learning.

Serhii Kushch *– Ph.D. in Electronics, ISOC Member, Associated Professor, Faculty of Information Security, Institute of Physics and Technology, National Technical University of Ukraine "Kyiv Polytechnic Institute"; Postal Code 03056, Prospect Peremohy, 37, Kyiv, Ukraine; e-mail: skushch1@gmail.com.*

Major Fields of Scientific Research: Wireless Communication System, Ultra High Frequency Communication Systems, Advanced Signal Processing.

TABLE OF CONTENTS