APPLICATION OF INFORMATION TECHNOLOGIES TO MONITOR DAILY PHYSICAL ACTIVITIES

Vincas Benevicius¹, Vytautas Ostasevicius¹, Mantas Venslauskas²

¹Kaunas University of Technology, Institute of High-tech Development, Studentų g. 65, Kaunas, Lithuania, vincas.benevicius@gmail.com, vytautas.ostasevicius@ktu.lt
²JSC “Baltec CNC technologies”, Raudondvario pl. 148, Kaunas, Lithuania, mantas@baltec-cnc.com

Abstract: The paper presents a complex system for monitoring physiological data and physical activity levels in everyday life. System’s global model is presented; different parts are discussed in more depth. System’s features such as possible data sources, testing capabilities and support for different modules are presented and possible application scenario is given.

Keywords: Daily physical activities, complex system, physiological, biomechanical data.

1 Introduction

The technology that is available today provides new tools for monitoring biomechanical and physiological data during daily physical activities. Rapid development of MEMS technology allows more and more non-intrusive data acquisition ways [1] that allow enhancing the experience during daily physical activities or even sport workouts. The application of latest technologies and innovations can provide the ways for data management processes automation and data acquisition rate improvement [2] which can boost research outcome, provide more motivation for reaching targeted goals and give a control over processes that were left aside simply because there were no easy to use and relatively cheap tools.

2 Problem description

There are a number of “intelligent” sports equipment producers such as Polar [3], Suunto [4] or Nike [5] that targets the group of active people and skilled sportsmen. There are series of other high-end equipment producers that target very specific sports or intended to be used by high level athletes as are Sports Performance Indicators from GPSports [6]. There are also research targeted devices such as Actigraph [7], but all of them are not intended to be used in everyday sports, physical activity monitoring or automated testing. All these systems are still quite expensive for personal use, and team solutions put the price tag even higher. That is why the System for Monitoring And Refined Testing (abbreviated as S.M.A.R.T.) is being developed to touch the parts the others are missing:

• To bring innovations close to everyday life;
• To serve as a tool to increase personal motivation for physical activity;
• To serve as an easy, fast and reliable personal testing tool;
• To serve as a comfortable data acquisition tool during research activities;
• To increase interest in sports of the masses, especially the youth, and to indirectly serve as a prevention from obesity and cardio diseases.

3 System’s model

The whole system consists of three major parts with dedicated functionality:

• A personal wearable, portable device (or set of the wearable devices) for data acquisition,
automated testing, physical activity monitoring, workout control, energy balance control and scheduler support;

- A PC software to aid in acquired data upload and synchronization and to serve as a local monitoring tool in case of team type sessions;
- A website for data storage, visualization and analysis.

The system supports a range of data types – acceleration measurements up to ±6g on every axis with sampling frequency from 1Hz to 128Hz from up to 4 end points with wireless link (2.4GHz). Acceleration module highly profited from the analysis of accelerometer model under dynamic excitation [8]. Other data types – angular velocity measurements up to 300 degree/second with sampling frequency from 1Hz to 128Hz, RR intervals using Polar or similar 5kHz coded heart rate belts and one derivative ECG using the chest belt designed by Biomedical Engineering Institute [9] with wireless link (also 2.4GHz).

All data flow in the system is presented in the following figure.

![Data flow in the S.M.A.R.T. system](image)

S.M.A.R.T. provides full range of data types for acquisition. Each data type is available either directly via base module (small handheld device with LCD screen and data storage) or via 2.4GHz wireless body network end points (small mobile devices with measuring and data transmission capabilities). A research [10] to optimize wireless body network parameters in similar applications was conducted; results and recommendations were used to design this particular wireless network in the system. Specific scalable star type network protocol was created to fulfill the needs of the required data rates.

### 3.1 Device

The device serves as a centre of wireless body network subsystem. It is designed to be a tool for data acquisition from the inner sensors and the other modules of the wireless body network, initial data processing and operation configuration and provides a simple interface for the end user via graphical LCD and few buttons.

**Primary data analysis** is done in the device to count steps while walking / running based on the second order maxima with a penalty function; to calculate energy expenditure (based on research that was carried out during another project “Development of Methods and Equipment for Human Active Motion Monitoring” [11]); to calculate walking and running speed; to calculate activity index (index describing the “weight” of the motion intensity); etc.

Another important function of the device is automated testing of various physiological conditions. **Automated testing** with guidance includes Ruffe test for the evaluation of the functional state of cardiovascular system, orthostatic heart rate test to monitor state of health; VO2 step test for the evaluation of maximum oxygen uptake VO2max, test to measure jump height from a static standing position and body mass index according to the personal data. The results for each test are immediately presented on the screen of the device. More tests will be integrated in the future or upon commercial request.
The device also **supports the scheduler**. Scheduler is designed the way so it could be the aid during workout process. Workout session or session sets are assembled using web portal tools, later these sets are then downloaded to the device and are run in the device to control the duration and the sequence of the exercise as well as resting periods.

A set of **workout tools** are integrated into the device to aid during workout process. They include the possibility to add time marks with a push of a button, metronome with user defined interval for periodical signal generation and pulse control tool to inform the user when person’s heartbeat goes off the targeted zone that is set using integrated tests or web server tools.

Diet control is very important aspect both in everyday life and sports, especially when reaching for specific goals like weight loss / gain or sugar control for diabetes, etc. **Diet control** is also integrated into the system. A personalized food menu can be set up with different approach levels (favorites, complex meals, etc) using web server tools, then assembled menu can be downloaded to the device with a help of PC software to allow food input selection in the device. Such functionality allows much strict control over overall process as energy balance can be provided anytime (with condition that the device is worn all the time for accurate energy expenditure calculation).

### 3.2 PC Software

The main task for the PC application is to synchronize data between the device and web server. The application allows:

- Synchronize personal, team and scheduler data between the server and device;
- Upload during workout or daily use gathered data to the server;
- Update firmware in the device for added functionality and bug fixes.

The application is very user friendly, is made wizard oriented way, so no additional fuzz is added for the user. The screenshot of the application can be seen below.

![Figure 3. S.M.A.R.T. PC Software screenshot](image)

### 3.3 WEB Portal

Another major part is web server application with data visualization and analysis tools. All is available online for easy access and rich end user experience. The portal serves as central data storage where every data from every S.M.A.R.T. device is safely kept. Using tools provided via portal this data can be analyzed, viewed and downloaded in a text format for further analysis using some external analysis tools.

A few screen shots from calendar and day views are presented below.

![Figure 4. Calendar and day views in the S.M.A.R.T. website](image)

Calendar view is provided for easy overview of the events – days containing any are marked with specific icons. Navigation with drop down list and navigation buttons is easy to use. Each day title in the calendars is linked to a day information view that is presented below. Day view provides in depth view on the events for a given day. This includes profile changes, workout results, test results, etc. Everything is provided with full text explanations.
The whole system is designed the way that nothing that is uploaded would ever be deleted. Such approach allows collecting population data and allows getting the statistical results not only from personal or coach point of view, but also the statistical results for the population itself. Such statistical data can further be used to conclude on the state of the population let’s say from the physical activity level point of view and to compare with the data of other populations.

3.4 S.M.A.R.T. features and capabilities

S.M.A.R.T. is designed to be used in daily life to monitor personal physical activity levels, to control energy balance by calculating in and out calories, to aid in workout processes for heart rate control and control of the workout itself and to aid in scientific researches as a data acquisition tool.

From what is said, S.M.A.R.T. has these key features:

- Number of data types available for acquisition with configurable sampling rates;
- Latest MEMS technology, miniaturization and use of 2.4GH allows wearable devices to be small and unnoticeable with capability to transmit data wirelessly;
- No deletion policy allows gathering data and making statistical analysis not only using personal but also whole population data;
- Everything is accessible online from any point where the internet connection is present;
- Easy integrated tests allow quick evaluation of personal fitness level;
- A tool for energy balance calculation;
- Heart rate and workout process control;

Given these features the system can be applied in daily life as a physical activity monitoring solution and help people to stay fit to avoid cardio diseases and prevent obesity. The system can be used in sports as an easy fitness level testing tool, heart rate and workout sequence control tool, or it can be used as a data acquisition tool for various researches.

4 Conclusions

There is a number of equipment capable of measuring biomechanical and / or physiological data, but most of the time such equipment is either expensive, targeted for very specific audience or is difficult to use in everyday life. System for Monitoring And Refined Testing covers most of mentioned deficiencies by providing a complex solution that is easy to use, does not have high price tag and has the number of attractive and useful features. The potential applications of the described system is huge: either its children / adults physical activity level monitoring and control, workout control for different types of sports, including fitness clubs, or even data acquisition for research purposes.

References


ENERGY CONSUMPTION DATA ACQUISITION AND DISPLAY SYSTEM DEVELOPMENT

Gintautas Mumgaudis1, Violeta Bulbenkienė2, Audrius Kuprinavičius3, Evaldas Zulkas3, Aleksej Pecko2, Aleksandr Sokolov3

1UAB Koris, Danes str. 6, Klaipėda, Lithuania, gm@koris.lt
2Klaipėda University, Department of Informatics Engineering, Bijunu str. 17, Klaipėda, Lithuania, bulbenkienė@gmail.com, aleksej.pecko@gmail.com
3Klaipėda University, Department of Informatics, H. Manto str. 84, Klaipėda, Lithuania, kuprinavičius.a@gmail.com, zulkas.e@gmail.com, sokolov.le@gmail.com

Abstract. This work describes the creation of a system that acquires, stores, and presents energy consumption data. This system is designed to visually present the user with energy consumption information and energy consumption trends. The system can accept data in any format using standard data communication protocols. This work shows how the data is acquired from the data collector using a TCP/IP protocol, converted to a format that is accepted by the database using a background service, stored in a secured database, and displayed to the user via web interface. A block schema of this system and code fragments are presented in this paper.

Keywords: data acquisition, data storage, monitoring system, visual data presentation, data conversion methods.

1 Introduction

Rising energy consumption and energy prices increase expenditures. In order to efficiently use electrical energy, a monitoring system that provides energy consumption data collection and mapping is necessary [2-4, 7]. Currently such systems can only be afforded by large businesses. Similar systems that are available for ordinary household users don’t have the functionality required to store large amount of data and because of this fact the results are not reliable, also such systems are expensive and the payback period takes a very long time. Our goal was to create electrical energy consumption data collection, transmission and storage system for households. System structure development was dependant on designing and operating costs. Monitoring results must be easily accessible to the user not only by the computer and the Internet, but also by the alternative means of communications such as fax, email and sms messages. Developed system's popularity and success depends on the cost of the system, its installation and maintenance simplicity and easy way of tracking the data. The system must automatically track and collect data of energy consumption. This work provides an analysis of energy consumption data storage system with web interface where user can monitor energy consumption trends and an analysis of a database interface with web development.

2 Data collection, transmission and storage

The system combines several important elements which carry out their assigned functions such as data collection, transfer and storage. Principal scheme of the system operation is presented in Figure 1. This paper examines the electrical energy data reading mode, which does not require any direct intervention in the electricity grid infrastructure.

Electricity consumption scanning module collects information through the device that relays input data to the consumer in his chosen form. Data collection begins with the physical energy consumption data read. Sensors are located right at the entry of the energy resource to the consumer house or on the power meters. Sensors can also be placed inside the consumer’s household to measure devices separately. After reading the data, it is transferred to primary storage. From primary storage, data is sent to a wireless module, which broadcasts data to a concentrator (Figure 1. stage 1). Energy consumption data is transferred with service information, which allows addressing and managing terminal nodes [6]. Information in the concentrator is provided to the CPU, which stores data in its memory (Figure 1. stage 2). When certain amount of data is collected it is forwarded to the hub via the Internet to a special server that manages data transfer (Figure 1. stage 3). The server processes the data and transmits it to the database (Figure 1. stage 4). The user can see all the historical information on the database server over the Internet connection (Figure 1. stage 5). The system is left with the possibility to transmit data not only via Ethernet protocol. Other transmission methods are possible via standardized data transmission protocols such as GSM, IEEE 802.14.5 (ZigBee Pro), IEEE 802.11 (Wi-Fi), RS485 and other wire and wireless data transmission protocols. System operation is automatic and does not require human action for its full performance. The system components can operate independently for a short time without communication with other parts of the system.
without data loss [7]. Autonomous operating time depends on the data reading frequency and the concentrator RAM.

![System operation principal scheme](image)

**Figure 1. System operation principal scheme**

### 2.1 Concentrator interface with the server

Concentrator kernel is a real-time task management system. It makes task queue in which tasks are carried out after the scheduled time according to priority. Tasks are carried out asynchronously. Upon receipt of a higher priority task, active task execution is interrupted and continued when there are no higher priority tasks. Microcontroller is used in conjunction with the network controller as TCP client and via Ethernet network can transmit data to the server [3]. Energy resource consumption data must be sent in a sequence from the hub in pre-defined structure. Concentrator packet structure is shown in Figure 2.

![Concentrator packet](image)

**Figure 2. Concentrator packet**

TCP/IP packet:

1) CONCENTRATOR_ID – concentrator identification number, 2 bytes.

2) SENSOR_ID – sensor identification number, 2 bytes.

3) WIZNET_CC – byte size command code (possible codes: „T” – time, „D” – data). If necessary it is possible to use up to 256 different codes to identify commands.

4) WIZNET_PARAMS_COUNT – the amount of data transmitted. The maximum amount of data sent is 255. Unless MAX_PACKET_SIZE indicates otherwise.

5) WIZNET_PARAMS – an array of data to be transmitted.

### 2.2 Receiving data

Consumer doesn’t need to worry about running databases or collecting data from the concentrator. This is done by the service provider which manages data collection and storage. User is provided with login details which he uses to login to a web page where he can monitor his energy resources consumption trends.

Special-purpose computer that functions as the TCP server collects data sent by the concentrator and forwards it to a remote database. To develop a background TCP server program it was decided to use C#.NET framework because it has rich libraries for network socket programming and is easy and fast to implement. Making a smaller, cheaper, and more efficient product is one of the main criteria’s, therefore more research needs to be done in evaluating the effect on the mentioned factors by switching to other technologies rather than C#.NET.
The background program must be running at all times in a closed cycle and accept data from the concentrator. Upon receipt of the TCP message from the concentrator, the server has to handle it in accordance with energy consumption data structure. System.NetSockets library makes working with network sockets possible and System.Threading allows us to use threads in a C# mechanism.

TCPListenerThreadRoutine() forwards the work of transferring data to another function that processes data received: tcpProcessMessage(byte[] message, TcpClient client). The function has two parameters: the transfer TcpClient client and message of the data array. Function that processes data sends the system time to the concentrator, processes energy consumption data and sends it to a remote database.

Data between the concentrator and the database can be forwarded not only via Ethernet protocol. However, each variant of the protocol must have a separate special purpose background server program to receive, process, and transmit data to the database. To implement the link via alternative protocols, server requires special software and hardware that supports the protocol integration (GSM transceiver, an IEEE 802.11 receiver). In this case, scheme in Figure 1 could be realized without an Ethernet router.

2.3 Data processing and storing procedure into the database

Data accepted by the background service, must be formed in special format required by the database.

Data sent from the concentrator in TCP/IP packets is prepared for hosting in the database. With this data no further action is needed, because it already has passed initial filtering and data processing. Background service has data filtration algorithms that can be used if the sensor system microcontroller does not provide data filtering.

Database information update has been realized in the algorithm:

tcpProcessMessage(byte[] message, TcpClient client) accepts a TCP/IP packet, processes it and call a recursive database update function updateSqlFunc:

```
updateSqlFunc(int meterID, Datetime dataStartDate, Int16[] dataArray)
{
    //Data Check:
    checkDataIntegrity(Int16[] dataArray)
    //Synchronization with the database
    requestSqlDataCheck(int meterID, Datetime dataStartDate)
    //Preparation of SQL procedures
    buildSqlQueries(int meterID, Datetime dataStartDate, Int16[] dataArray)
    //Data transfer database
    executeSqlQueries(string[] queries)
}
```

The database can accept data from other similar systems, but they must submit data in a special form.

2.4 Data extraction from the database

Data review is realized via HTTP protocol. Website is developed using ASP.NET technology. The site is interactive and a dynamically updates real-time energy consumption data that just came to the database, thus improving the user's experience using the system. One data point represents the cumulative values of consumed energy per minute. This data density allows accurate visual representation on energy consumption. Data is continuously collected 24 hours a day all year round. The user can save the displayed page graphics by the selected time interval [1].

In future plans, server will have a mail server and GSM interface. User will be able to send selected settings to the mail server and a reply will be sent back to the user with information about energy consumption. GSM users will be able to obtain a report by SMS to check whether the energy resources were used recently.

User can download energy consumption data from the site in CSV and XML format. In the future it is expected to extend data export format selection.

3 Conclusions

Developed electrical energy monitoring system consists of three main parts: energy consumption data metering device, transmission system designed to transmit data to various data storage systems and a system which graphically displays the received data and generated data reports. Automatically operating monitoring system does not require external human interaction; therefore the user doesn’t need to manually record the meter readings. Data can be quickly reviewed, understood and compared. This saves considerable time and the user can
deal with big amounts of consumed energy data. The results obtained will contribute to more efficient use of energy resources and reduce costs.

Developed interactive Web interface is capable of visualizing real-time energy consumption data and it is used to review energy consumption data history. In the future it might be possible to access reports via email or SMS.

Concentrator and database interface is created using TCP/IP protocol. Server is running a background service that transforms the TCP/IP packets to the energy consumption information. Analysis has shown that it is possible to dispense without the background service application and send the data directly from the concentrator to the database. To use another data transmission protocol it is necessary to prepare the necessary equipment and background services.

In the future this system can be used not only for electricity, but for other sources of energy like gas and water usage monitoring. This can be done by using already developed data acquisition module and by replacing the sensor and data processing software code.

References


SYSTEM FOR MONITORING A VEHICLE’S DYNAMICS USING AN ACCELEROMETER AND FUEL INJECTION SYSTEM SIGNALS TO ASSESS DRIVING ECONOMY

Mantas Venslauskas¹, Vincas Benevicius², Gintaras Rimsa¹, Vytautas Ostasevicius³

¹JSC “Baltec CNC technologies”, Raudondvario pl. 148, Kaunas, Lithuania
mantas@baltec-cnc.com, gintaras@baltec-cnc.com
²Kaunas University of Technology, Institute of High-tech Development, Studentų g. 65, Kaunas, Lithuania, vincas.benevicius@gmail.com, vytautas.ostasevicius@ktu.lt

Abstract. System that is dedicated to monitor and analyze vehicle’s dynamics is presented in this article. Data flow from accelerometer and signals from injection system provides a needed data to assess driving economy. System would solve incorrect car’s driving problem in companies’ fleets, also reduce CO₂ emission amount and teaches drive more economically herewith system is dedicated for a personal use getting indications about driving manners. Broadly adoptable system is designed with ability to expand its features regarding to up to date technologies and market demand without significant changes.

Keywords: driver’s behavior, reduce CO₂, ecodriving, accelerometer, injection system.

1 Introduction

The whole world is trying to find a solution to reduce emissions. Car’s manufactures are squeezed in borders that are narrowing constantly. However, it is stated that the abilities of automotive industries is almost already reached, therefore it is needed to get bigger impact from drivers of the vehicles. The cars that are currently produced have a particular economy but until the 2012 the amount of emission should be reduced by 120 g CO₂/km. While taking advantage of the technologies that evolve constantly, integrated modular sensors and correctly interpreting the data and linking gathered data with additional respective branches of the system, it is possible to get wide range spectrum of data to anticipate the monitoring of multidisciplinary to get ability to analyze and to define the driver’s behavior in the same way enabling to teach/learn economical driving.

2 The definition of the problem

When talking about equipment and systems that are made to monitor vehicles there are similar trends in Lithuania’s and whole world’s markets, however, different in many aspects. The markets are fulfilled with products with similar functionality that has many shortages in many cases. Systems that are spread in Lithuania’s market are oriented at companies’ fleets so it is difficult to use them for personal needs. The products from abroad markets’ are oriented both at companies’ fleets and personal consumers but have limited ability to link with various vehicles because of connection to get necessary data through OBD II interface. The data which is possible to get from an on-board computer is informative but insufficient to assess driver’s behavior on the road at instantaneous situations. For this and other actual seen problems it was decided to implement EcoDrive system which would let to informatively define the manners of driver’s behavior, to perform a monitoring of a huge flow of the vehicles, give informative reports of the trips which would allow get tangible benefits. The data from the system with complete functionality would be informative and enable user to analyze the performance of employs that are using vehicles of companies’ fleet herewith gathering and supplying data of precise usage of fuel, the overcame kilometers, routes on the map and the economy of the nature of driving.

3 System’s model

System consists of three main parts:

- Vehicle – a part of the system where EcoDrive device is integrated. The device gathers data and processes various parameters (eg. dynamics of the vehicle). Also there is an in-car installed system to monitor instantaneous fuel consumption which is based on electrical impulses of fuel injection system that controls nozzles.
- EcoDrive device – equipment to gather, store and process data with interface to transfer it.
- Web site – the interface between consumer and remote server which enables to monitor and analyze data of the routes, monitor progress and regress all the time. Also place for discussions on related topics between system’s users is also foreseen there like route reports, analysis of statistical driver’s assessment.
The assessment of driver’s manner is based on EcoDrive system’s accelerometer data and instantaneous fuel consumptions that are available when analyzing injection system’s control signals.

Figure 1. EcoDrive system’s modular scheme

The EcoDrive system’s modules and data flow between them is showed at Pic. 2. Similar data could be gathered in other ways but the chosen one is the most precise and optimal with regard to needs of system’s functionality. The integration of additional models that would expand the functionality is foreseen after the pilot tests and recognition of consumers’ needs. It is also planned to install wireless module for data transmission to provide the ability easily work with the data.

Gathered data is sent to the remote server for further analysis and to provide it to the user with an attractive form. The safety is guaranteed when allowing only look through the data without ability to delete them. All data from all users that is stored at one server could be used for general statistical analysis and comparison between other system’s users.

Figure 2. Data flow between EcoDrive system’s components
**EcoDrive system’s features and advantages**

The monitoring of driver’s behavior on the road and the outlet of results and recommendations while using two types of indications (audio and video); at the same time storing all the data inside the device for further analysis;

- The precise monitoring of fuel consumptions;
- The precise monitoring of the routes on the map (after installing GPS module while using open code maps);
- The real time data transmission (after installing GSM / GPRS module);
- The analysis and review of data is possible anywhere anytime just having a connection to the internet;
- The ability to store and calculate exploitations costs (using WEB site);
- To control the traffic, economy and driving manners of employs’ vehicles like the work time of engine, the blow away of fuel and etc.;
- The ability to apply system as prevention from theft (after installing GSM / GPRS module);

**The application of the EcoDrive system**

- For individual user – to cultivate economical driving habits;
- For company’s fleet – to the control and cultivation of economical driving habits:
  - Management companies;
  - Transport from service sector;
  - Companies that makes a passenger’s transportation;
  - Companies that makes a cargo’s transportation and etc.
- Driving schools:
  - To learn ecodriving at primary courses;
  - Making additional specialized courses for ecodriving learning;
  - To apply system with regard to “eCall” provisions.

**4 Conclusions**

With reference to the concurrence environment analysis it could be said that there is a wide range of similar products at the market, but there is no one with exact functionality. Only the little part of them has ability to provide information about driving manners either they are limited to adopt. The application of the developing EcoDrive system is very wide starting from individual driver and ending with companies’ fleets when installing system for full vehicles control. The application of the system could be extended when synchronizing it with “eCall” system because all technical capabilities provide ability to meet all requirements and goals.

**References**


AN ULTRASONIC TRACKING METHOD FOR AUGMENTED REALITY

Eduardas Lukosiunas, Violeta Bulbenkienė, Arunas Andziulis,
Linas Pasviestis, Edgaras Artemciukas
Klaipėda University, Department of Informatics Engineering, Bijunu str. 17, Klaipėda, Lithuania, eduardas.lukosiunas@gmail.com, bulbenkienė@gmail.com, arunas@i.ku.lt, linas.pasviestis@gmail.com, edgaras.a@balticum-tv.lt

Abstract. Augmented reality (AR) is an emerging technology that is able to integrate virtual objects onto the real world view in order to develop new generation human-computer interaction interfaces. Nowadays position tracking systems used for virtual object augmentation are based on image processing technology that identifies real world objects with special markers and visual quality parameters. In order to assure reliable positioning for objects in the virtual space, new advanced solutions are required, therefore the real world object positioning system based on ultrasonic tracking method is introduced. The prototype of the system was developed and an object positioning tests were made to verify the efficiency for the new solution. Acquired results were compared to the other tracking systems in order to demonstrate the increase of accuracy and quality parameters.

Keywords: augmented reality, ultrasonic tracking, object positioning, virtual objects.

1 Introduction

New technology was derived from virtual reality (VR) called augmented reality (AR) in the recent decade. Computer generated 3D objects are blended onto a real world view in AR environment in order to enhance user’s interaction with the real world. An AR system basically consists of two main activities: tracking, where the orientation and position calculation for the camera is based on a real world objects and registration, where virtual objects are rendered onto the user’s view, based on the values that are calculated from the tracking process. The computer performance has increased in the last decade, therefore the 3D object development and navigation performance in virtual environment has also increased. Architectural and engineering designs are often developed using virtual methods, in fact 3D models become more complex and photo realistic over the past years. Additional information about the environment and virtual objects can be stored and rendered as a visual information on top of the real world view [7]. Observer is able to see the real world image augmented with virtual objects through HMD (Head Mounted Display), thereby AR provides detailed information about environment, possibility of smarter navigation and increase of intuition while performing tasks [5]. However, there are some issues related to the mobility, ergonomics, limited field of view and insecurity when moving around. Position of the virtual world object must be synchronized with the position in the real world, therefore computation of coordinates must be very accurate. In order to track objects smoothly, position of the camera must be calculated for each frame. Currently, such computations are mainly performed by using visual tracking method. While camera is moving, aspect of object is changing. According to these visual changes computation of an object’s position is performed [1]. Nevertheless, such complex measurements are suitable only for immobile or slowly moving and fully visible objects, otherwise tracking of an object is canceled. This paper suggests objects tracking method based on ultrasonic sensors that is capable of camera position estimation in order to generate AR (discussed more in section 4), that was also successfully applied for a four-rotor flying robot’s position estimation in a limited indoor area [2].

2 Related work

AR is an emerging technology that is becoming important in these directions: medicine, robotics, engineering, etc. The main difference between AR and VR is that AR eliminates the need to model the entire environment as it supplements the real-world instead of replacing it.

Recent research in tracking methods show positive results achieved by integrating magnetic tracking and trajectory computation [3]. System demonstrates the ability to compute the trajectory accurately for dynamically moving objects. Another solution, ARCam system, goes even further, introducing tracking subsystem that provides high accuracy and precision in various viewing distances and mobility while covering a large range [8]. Satellite positioning systems like GPS take AR outdoors [6], but are still miserable in accuracy. In fact camera parameters, while using vision based tracking technologies, are very important in order to match real and virtual world coordinates accurately. Accurate calibration is especially important with 3D display systems, because when rendering virtual objects above real environment, they have to match the perspective and depth of the real scene. User would certainly get very disoriented if virtual objects would be floating in front of real objects. A necessity of AR systems also is the latency. Because some processing has to be done in order to dynamically supplement reality with virtuality. If the decrease in details is not plausible answer to decreased latency, it has to be tackled with some other way. Prediction, pipelining and other tricks could be used in
microcontrollers to reduce latency. Autonomous ultrasonic indoor tracking system (AUITS) processes signal acquisition and conducts position computation [9], but the results show lack of accuracy.

To sum it up, all methods lack in some area: AUITS, GPS and magnetic trackers systems are still not so accurate, mechanical trackers are cumbersome and vision-based trackers are computationally problematic.

3 Mathematical procedures

This section covers the procedure of computing object’s position information out of gathered distance measurements between receivers and transmitter. The technical details of tracking system prototype are discussed in Section 4.

3.1 Distance calculation

We assume to start with a set of \( n \) tuples \( T_i \), each consisting of a distance \( d_i \) to a reference point with a known position and the coordinates of this point \( (x_i, y_i, z_i) \) :

\[
T_i = (d_i, \bar{x}_i); \quad \bar{x}_i = (x_i, y_i, z_i)^T; \quad i \in [1, n]
\]  

In order to determine the position of transmitter in space \( \hat{x} = (x, y, z)^T \) where distances are known between the transmitter and three (or more) receivers, the trilateration problem should be solved by using a closed mathematical expression:

\[
(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2 = d_i^2; \quad i \in [1, n]
\]  

The distance traveled by sound depends on its speed, which has a direct dependence on the environment temperature. The distance between a transmitter and a receiver is calculated:

\[
d = t_v \cdot v + d_n,
\]

where \( d \) - distance between transmitter and receiver [m]; \( t_v \) - travel time of sound [s]; \( v \) - speed of sound in certain temperature [m/s]; \( d_n \) - minimal distance, that receiver can locate [m].

System prototype testing was made in an environment where temperature is equal 20°C, consequently we assume that the speed of sound in this temperature is equal 343 m/s.

3.2 Position calculations

Coordinates calculation of an object is performed according to three projection planes as shown in Figure 1.

![Figure 1. Position calculation of an object](image)
Since positions of the receivers I1 \((x_1, y_1)\), I2 \((x_2, y_2)\), I3 \((x_3, y_3)\) and I4 \((x_4, y_4)\) are always static, hereof coordinates are known and constant. In the first place, data of I1 and I2 receivers are processed accordingly to the first projection plane \(S_1\). The presented equation system (4) is based on (2) formula in order to calculate \(x\) value by using (6) formula:

\[
\begin{align*}
(x_i - x)^2 + y^2 &= d_i^2, \\
(x_i + x)^2 + y^2 &= d_{i+2}^2, \\
(x_i - x)^2 + (x_{i+2} - x)^2 &= d_i^2 - d_{i+2}^2, \\
y_{n1} &= \sqrt{d_i^2 - (x_i - x)^2}.
\end{align*}
\]

Secondly, data of I3 and I4 receivers are processed accordingly to the second projection plane \(S_2\). \(y_{n2}\) value is calculated by using (6) formula. Then, \(y_{n1}\) and \(y_{n2}\) variables are used as distances to calculate \(y\) and \(z\) values accordingly to the third projection plane \(S_3\). Therefore all three necessary coordinates \((x, y, z)\) are calculated.

### 4 Technical prototype overview of a new tracking method

The prototype of tracking system is using four 400SR120 ultrasonic receivers and one 400ST120 ultrasonic transmitter. Four receivers are located on different corners of the room, three to four meters apart from each other, allowing up to four measurements of a distance at the same time. Ultrasonic transmitter that is capable of generating 40 kHz ultrasound wavelengths, forms and sends signals to all receivers in sight at a fixed time of 40 ms, but receivers get the signals at different moments (depending on transmitter position). Transmitted signal consists of four ultrasound waves that are generated in 100 \(\mu\)s. According to this time theoretical \(d_n\) constant (3) is equal to 3.43 cm, which marks the minimum distance that can be measured between the transmitter and receiver. Practical accuracy of distance measurement is estimated to be approx. 0.515 mm, since travel distance of sound during 1 \(\mu\)s is equal to 0.343 mm and remaining time ~0.5 \(\mu\)s is spent for algorithm execution. In order to increase accuracy of distance measurement, system should use greater number of receivers. The prototype of tracking system is shown below (Figure 2).

![Figure 2. Prototype of the system](image)

In order to estimate 3D object coordinates, microcontroller is used to process distance data. The stream of data packets consisting of estimated object coordinates are sent to the computer, where special software (with DirectX support) renders an object in the 3D virtual space. Communication speed between microcontroller and computer on used RS232 protocol is equal to 9600 bps. Since data is packed in the 40 bit structures (Table 1),
this provides theoretical transfer rate of 240 packets/sec. This satisfies the condition that object should be updated more than 25 times per second in order to ensure smooth rendering of an object in the 3D virtual space.

### Table 1. Structure of communication packet

<table>
<thead>
<tr>
<th>Element</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>Reserved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>12 bit</td>
<td>12 bit</td>
<td>12 bit</td>
<td>4 bit</td>
</tr>
</tbody>
</table>

The experiments were carried out to obtain bias with different results of distances, while changing the time \( t_b \) in the algorithm: in the first attempt \( 490 < t_b \), or otherwise, the distance is registered after 490 s. Receiver accepts signal with longer delay when deviation reaches 45 degree, therefore the estimated distance is 6mm longer than the theoretical. During the second attempt the time in the algorithm was changed to \( 990 < t_b \). The distance we discovered was twice longer and bias increased to 12 mm. It means that relative bias depends on measured distance and can be calculated as a percentage:

\[
\varepsilon = \left( \frac{d_r - d_t}{d_t} \right) \cdot 100 \tag{9}
\]

where \( \varepsilon \) - relative bias [%]; \( d_r \) - real distance [m]; \( d_t \) - theoretical distance [m].

Calculated relative bias was 3.5 % so if measured distance is equal to 1 m, then the bias is 3.5 cm. In order to decrease relative bias to 1% it is necessary to use three receivers at one point.

### 5 Conclusions/Future works

The main advantage of using ultrasonic tracking method is that this method provides fast position computation of an object and more reliable system compared to the existing methods, based on visual tracking. Ultrasonic tracking method justifies the requirement of accuracy that is necessary to get high quality AR views. Tests demonstrated that the system can only detect objects within a maximum distance of 13 meters. For this reason, the system cannot be applied in an outdoor environment, however it can be well suited in a room, classroom or car cabin space. This system can be integrated with GPS [4] to use it in outdoor environment and obtain better results.

Since the AR requires visualization through HMD, further research will be presented on limited field of view, lack of resolution and insecurity when moving around. Prototype will also be improved by integrating IMU (Inertial Measurement Unit) devices, in order to provide the feature to estimate pitch, yaw and roll parameters.

### References


STANDARTIZATION IN DESIGNING RESEARCH AND APPLICATION INFRASTRUCTURE FOR LITHUANIAN LANGUAGE TECHNOLOGIES

Vladislav Fomin¹, Daiva Vitkute-Adzgauskiene¹, Ruta Marcinkeviciene², Andrius Utka²

¹Vytautas Magnus University, Department of Applied Informatics, Vileikos 8, Kaunas, Lithuania, v.fomin@if.vdu.lt, d.vitkute@if.vdu.lt
²Vytautas Magnus University, Centre of Computational Linguistics, Donelaicio 52, Kaunas, Lithuania, r.marcinkeviciene@hmf.vdu.lt, a.utka@hmf.vdu.lt

Abstract. The paper overviews information infrastructure standardization problems, as well as particular problems of Human Language Technologies (HLT) in Lithuania. The main focus of the paper is on the standardization of HLT infrastructure in Lithuania, whereby emerging European infrastructures are introduced, main language resources and their standards are analysed, guidelines to the roadmap for Lithuanian HLT development is presented. A pan-European infrastructure CLARIN is named as the most promising of emerging infrastructures, its recommendations of standards for resource description, resource access and metadata harvesting are summarized. A pilot experiment of setting up a CLARIN metadata providing centre is presented.

Keywords: standardization, research infrastructure, interoperability, language technologies, CLARIN.

1 Introduction

Human Language Technology (HLT) development in Lithuania as in many other countries started with the development of corpora resources. Corpus of Contemporary Lithuanian Language CCLL, Corpus Academicum Lithuanicum CorALit, Corpus of Spoken Lithuanian language, Corpus of Lithuanian Dialects and Database of Old Lithuanian Writings, several parallel corpora are currently available, as well as different computerized bilingual and multilingual dictionaries, a number of smaller or larger lexical databases and other language resources. In line with resource compilation, various language technology tools are available or in design process by different institutions, e.g. concordancers, collocation and terminology extraction tools, spellcheckers, morphological analysers and annotators, lemmatizers and corpus aligners, accentuation tools, translation systems, syntactic and semantic analysers and taggers, etc. Research, as well as compilation of resources and tool design are also done in speech signal processing and synthesis. Results include speech corpora, (corpus of spontaneous speech, corpus of digit sequences, radio news corpus, etc.), also, speech recognition prototype for recognizing voice commands, prototypes for continuous speech recognition, text-to-speech synthesizer, etc. [1].

When analyzing the overall structural picture of Lithuanian HLT resources, it can be concluded that foundations, i.e. a collection of text and speech corpora and other resources exists, though its expansion is still a task. There are too few language technology tools available, especially those belonging to the advanced and general application categories. Existing tools are mainly oriented towards researchers, few on-line tools are available for end-users or application programmers. Corpora and other resources are designed by different institutions and available only via specialized, individual access tools, or even unaccessible. In many cases, standards, such as TEI P5 or simple XML, are underused.

A central tenet of the knowledge society is about maximally even distribution of knowledge among the citizens and businesses, and not just about availability of information on the Internet [1]. What is needed to get from the current state of the art in HLT in Lithuania (disparate tools and corpora) to the common service paradigm (inter-operable resources and automated services)? The information infrastructure is often defined as comprehensive, interconnected, and integrated modules of information systems [2,3]. Will the task of integrating these disparate HLT modules into a service-providing infrastructure be a straightforward or a daunting one? To answer these questions, we must understand the driving forces behind the infrastructural development in general, and the specifics of the Lithuanian HLT in particular.

2 The information infrastructure

In his review of the history of computers as infrastructure [4], Paul Edwards writes that from its beginnings in the ARPANET of the late 1960s, the most fundamental principle of the Internet has been to assume heterogeneity in the networks being linked together. The basis for the massive interconnectivity which
we see in today’s globally interconnected networks, like the Internet, is a set of protocols, or software and hardware standards, developed over three decades by a community of hackers and computer professionals [4].

The critical role of standards as basis for interconnectedness of heterogeneous networks and formation of information infrastructures prompted interest of scholars representing different scholarly traditions, from economics, to history and social sciences. Although they all in one or another way acknowledge the special role of standards in the building of the information infrastructures, not much attention has been paid to the development of infrastructures outside boundaries of a single organization. Whereas a few works analyze development of inter-organizational information infrastructures, like the Internet [4,5,6,7], in most cases the focus is on corporate IT infrastructure development [8,9,10]. Although small, the number of works on infrastructures forms a solid ground for assessing complexity and versatility of the phenomenon, and is summarized as follows.

2.1 Requirements for information infrastructure development

Infrastructures evolve from different and relatively independent technologies that are meshed into a single overarching structure. To become integrated into information infrastructures, technologies must be harmonized in technological and socio-institutional terms, meaning that relationships between the elements have to be reinforced and transformed. In this, not only “the whole is more than the parts” [11], but a novel socio-technical configuration linking regulatory framework, communications infrastructure, user practices, maintenance networks, etc. [12] have to be formed. Standardization becomes a crucial element in such harmonization process, because of its ability to coordinate activities between and within diverse social groups to reconcile their varying interests while still working towards a single outcome [13,14]. If infrastructures are the connecting tissue of modernity [2], then standards are fibres of infrastructures. It is through aggregation of elements by the means of standards that results in formation of large and complex systems bringing about entirely new properties at each level of complexity [11].

From a technological viewpoint, information infrastructures demand standards that enable interconnectivity of multiple technologies, or “gateways” [15]. Creating “gateways” is a highly complex socio-technical task, which includes designing communication and technical interface standards, testing and adapting these to a wide range of different use situations, and ensuring that the standards are developed according to the procedures of recognized standardization bodies, if such exist [14,15].

Buildout of infrastructure requires alignment of diverse interests [16] of participating groups (government organizations, engineers, entrepreneurs, consumers, etc.) representing different (standardized) technologies or practices [13,15]. Standards provide a means for system builders and entrepreneurs to share their perspectives, and to gain understanding how the technological potential can be made to meet diverse ends [17]. In other words, infrastructural buildout requires inscribing and embedding of large socio-technical networks of developers, users, and government institutions. This can be achieved only through participating groups’ negotiation [18] over desired technical and economic properties of the infrastructure-in-the-making [19].

Negotiating which technological element or work practice must be chosen over its alternatives to become a part of infrastructure is often complicated due to the possible economic consequences of seemingly unimportant decisions. The reason for tough negotiation is the lesson that participating companies learned: competition between system standards leads often to a situations where “a winner takes all” due to strong positive network externalities and resulting increased returns [20]. A body or a firm, which succeeds in promoting its own technical standard in a new technological regime, typically receives large returns, whereas its competitors may be effectively locked out or provided only with residual market niches [21]. This motivates vendors to adopt protectionist policies for their products, forcing customers into lock-in to a specific product [22]. Such behaviors cause major obstacles in integrating communications networks, hardware, and disparate systems software into global information infrastructures [9].

Development of the Internet and its services has always been driven by the idea of interconnectivity. In recent years, a new trend of open standards has emerged to emphasize the democratic principle of the global internet and the requirements of the knowledge society.

The concept of open standards appears as a remedy against opportunistic strategies of private technology vendors. The presumption is that open standards lead to wider knowledge about technologies, more competition, which in turn, provides wider variety and lower prices for implementations of the standards and associated complementary products. Meanwhile, lower prices are expected to produce a positive feedback loop, further increasing demand [23].

For emerging infrastructure to gain a momentum, standards are both necessary and helpful in that they early limit the technical design space and help obtain a sufficiently fast implementation of a working design with

* Sometimes this “battle of the systems” can culminate with the invention of devices that make possible the interconnection between incompatible system [26].
functions favored by the dominant political-administrative coalitions in public administrations, and not those of functions, thus failing to support the principles of the knowledge society. “The primary beneficiaries have been little change to organizational structures, and often benefits of IT are not evenly distributed within organizational that an IT application

infrastructural technologies present as important. These requirements show that infrastructural development requires standards and standardization on multiple socio-techno-economic planes. But fulfilling any of these requirements separately does not guarantee successful growth of infrastructure.

The measure of success of regional and national informatization policies is bringing IT to everyday life of citizens and businesses, and improving effectiveness of communication between government administrations and citizens. Often, informatization programs are ineffective due to technical, cultural, political or other reasons [24,25]. As Kraemer and King [25] observe, the transformation of the administrative practices is not something that an IT application inevitably causes. Informatization of the public sector, for example, has brought relatively little change to organizational structures, and often benefits of IT are not evenly distributed within organizational functions, thus failing to support the principles of the knowledge society. “The primary beneficiaries have been functions favored by the dominant political-administrative coalitions in public administrations, and not those of technical elites, middle managers, clerical staff, or ordinary citizens” [25].

In developing the HLT services, we must get from the technical and political rhetoric of informatization of society to a concrete vision (or policy, roadmap) of how to overcome the existing fragmentation and enable user-centric service paradigm.

3 The Lithuanian HLT infrastructure

A modern Lithuanian HLT infrastructure is needed for the following user segments – researchers, designers of modern software tools, as well as private and business end-users. Researchers are interested in using aggregated linguistic resources and modern software tools for research in the computational and corpus linguistics, modern lexicography, etc. Software developers are aimed at the design of powerful tools for information and knowledge management, where language technologies, adapted to local language specifics, are especially important. Examples could include automatic translation systems, computerized media monitors, virtual assistant systems, semantic knowledge management, and semantic web services. End-users are looking towards applying language technology tools for the purpose of qualitative cooperation and collaboration on the web, for business information mining, market analysis, etc. End-user groups comprise social network users, media industry, marketing and information management officers in all kinds of companies, state institutions and organizations.

Despite the growing need for HLT resources and tools, so far the importance of the common HLT infrastructure was clearly undervalued. Only some attempts to solve this problem are visible. Design issues of the Lithuanian HLT infrastructure are addressed in the National Research Infrastructure (NRI) project e-Lingua (Electronic Lithuanian Language Resources). E-Lingua is recognized as one of priority NRI projects in the field of Social Sciences and Humanities (SSH) and included in the Lithuanian Roadmap of Research Infrastructures [27]. This Roadmap is built in alignment with the European Strategy Forum for Research Infrastructures (ESFRI), and emphasizes the necessity of international and, especially, European integration.

The institutional structure of e-Lingua NRI project opts for a consortium of institutions that would join their existing resources and jointly make new ones. It is planned, that it would consist of an a) administrative body providing data collections, methodological supervision and coordination inside the consortium of institutions and on the international level, specifically dealing with incorporation of Lithuanian language resources into corresponding ESFRI projects (e.g. CLARIN), b) associated bodies encompassing partner institutions which implement major SSH research projects in Lithuania.

Standardization is the first step in designing an open national technological infrastructure for wide access and joint use of Lithuanian HLT resources. E-Lingua is conceived as a common and highly interoperable virtual system of resources accessible via one main site, where resource aggregation would be done using federation-oriented infrastructure with linguistic resources stored both in centralized and in institution-specific digital repositories. E-Lingua defines, that resources should comply with internationally accepted standards (e.g. TEI P5 for the annotation of electronic resources). Unified resource access should be implemented using the standardized service-oriented mechanism for interoperable machine-to-machine interaction over the Internet. Open access to resources and tools is to be promoted for researchers and end-users in combination with flexible intellectual property protection mechanisms, applied by resource owners upon request.

However, design of the Lithuanian HLT infrastructure would require a detailed requirement specification for the standardization of HLT resources and tools, which is missing in e-Lingua. Specific recommendations for the alignment of recommended standard descriptions and overall architecture with European HLT infrastructure initiatives are missing. This paper addresses the aspects of this problem by
formulating suggestions for the design of a standardized HLT infrastructure for the Lithuanian language, possible solutions for ensuring its compatibility with main European networks and initiatives, selecting standard alternatives for different resource groups.

3.1 Interoperability framework

A common approach to overcoming interoperability problem under the conditions, where disparate technologies and standards exist in one service domain, is developing and agreeing upon an interoperability framework (IF).

A common IT architecture, or an interoperability framework, is defined as a set of standards and guidelines which describe the way in which organizations have agreed, or should agree, to interact with each other [28]. In other words, IF defines the use of common standards and guidelines and development of common methods for system design.

Having international integration in mind, a reasonable solution for the interoperability framework of a national HLT infrastructure would be to mirror or adapt an interoperability framework, defined by one of corresponding international (European) initiatives on HLT shared policies. However, a selection between several existing initiatives should be done in this case, the initiatives being CLARIN (http://www.clarin.eu), FlaReNet (http://www.flarenet.eu/), and META-SHARE (http://www.meta-net.eu/meta-share).

FlaReNet aims at developing a common vision and shared policy for the HLT area and fostering a European strategy for consolidating the sector. It maintains a FlaReNet Repository of Standards, Best Practices and Documentation with the aim of maintaining interoperability and reusability of resources [29]. The repository uses crowdsourcing mechanisms to file different documents, covering known guidelines, resource annotation schemes, tools, etc. In this way, FlaReNet only accumulates information on existing standards and best practices, but does not give explicit recommendations for balanced HLT infrastructure solutions.

META-SHARE is aimed at the development of a sustainable network of repositories of language data, tools and related services documented with high-quality metadata, aggregated in central inventories allowing for uniform search and access. Users will be able to discover and share resources across repositories, using supported metadata harvesting schemes. In order to ensure modularity and robustness, META-SHARE plans to follow a service-oriented architecture.

Though the aim of META-NET to develop an interoperable HLT-repository network is clear and reasonable, the project is in its starting phase, practically no technical recommendations are formulated. Therefore, it can hardly be used as a basis for the Lithuanian HLT infrastructure.

CLARIN as a pan-European initiative committed to establish an integrated and interoperable research infrastructure of language resources and its technology. It aims at lifting the current fragmentation, offering a stable, persistent, accessible and extendable HLT infrastructure [30]. CLARIN defines common HLT infrastructure as a network of centres of different types, representing different service levels. Technical requirements for the following CLARIN centre types is defined [31]:

1. Recognized Centres (Type R) offer resources and tools via standard web sites, but do not have funds to participate in the CLARIN infrastructure and cannot give commitment statements.
2. Metadata Providing Centres (Type C) offer machine readable metadata in a stable and persistent way allowing service providers to harvest their metadata and making them browsable, searchable and combinable.
3. Service Providing Centres (Type B) offer services that include the access to resources and tools via specified and CLARIN compliant interfaces in a stable and persistent way.
4. Infrastructure Centres (Type A) offer services that are relevant for the infrastructure as a whole and that need to be given at a high level of commitment. In contrast to Type B they offer services that are used by other centres.
5. External Centres (Type E) offer CLARIN relevant services but which are offered from non-CLARIN members.

It is a requirement for all centres that they adhere to the agreements defined for CLARIN centres. Requirements for a robust repository system, metadata descriptions, metadata harvesting via OAI-PMH, association of PIDs with their resources are defined. In some cases, e.g. for resource annotation metadata, alternative recommendations are given. CLARIN foresees the establishment of hierarchical structures for separate infrastructure elements, where national CLARIN contact points are to take responsibility for the national level of organization and quality, and the European CLARIN organization takes care of combining services at the European level.

Though CLARIN has not yet come up with an interoperability framework which can be followed straightforwardly (still too many standards), the scheme offered by CLARIN seems to be most reasonable for the
implementation of the Lithuanian HLT infrastructure. Therefore, the following recommendations and pilot implementations are based on CLARIN recommendations.

3.2 Developing a roadmap/policy

What could be the solution for Lithuanian HLT development, following the CLARIN recommendations, and taking into account the fact that choices between alternative standards have to be done for different infrastructure elements?

The literature postulates that in the presence of competing technology standards on the market, consumers make adoption decisions based on the perceived or anticipated value of a product or service [32,33,34]. In the eyes of the deciding consumer, compatibility (or interoperability) is of particular value. Due to the existence of network effects, the value of one standard is increasing with every new consumer joining the “network” of adopters. Consumers are more likely to opt for a product that is built upon a more popular standard, thus helping the market “tip” towards one standard [35]. In the case of HLT, however, the “old proven theories” on standards competition must be taken with caution. First, “consumers” are rather researchers than lay citizens. Thus, lay citizens’ choice, being “late adopters”, will be predetermined by the “early adopters” – the researchers actively involved in the shaping of the HLT infrastructure. Second, competition between HLT standards, which in many cases serve as digital converters, does not necessarily follow the well known logic of such standards contests as VHS vs. Beta or Mac vs. Windows.

In the Internet era, with hundreds of new standards competing, market (self-)regulation may fail (due to monopolization of one of the product offerings, or due to overly high fragmentation). In this case policy intervention is seen as justified or even required [36]. Seeing the fragmentation of the HLT infrastructure as a “system failure” [36], a policy/roadmap/interoperability framework must be developed and adopted.

Taking into account the CLARIN requirements, the following roadmap could be set for the Lithuanian HLT infrastructure:

1. Inventory of Lithuanian HLT resources and to be offered via common HLT infrastructure.
2. Selection of metadata standards for HLT resource annotation, covering all 3 main linguistic resource annotation levels (document level annotation, structure annotation and linguistic annotation).
3. Selection of metadata harvesting and resource access schemes.
4. Definition of Intellectual Property Rights (IPR) mechanisms for Lithuanian HLT resources.
5. Implementation of CLARIN Type C centres on institutional level, providing harvestable metadata.
6. Selection of standards and interfaces for HLT tool libraries, enabling interoperability and reuse. This step could just follow the CLARIN guidelines for service-oriented architecture, as there are not many alternatives here.
7. Implementation of CLARIN Type B centres on institutional level, providing access to HLT tools via CLARIN defined interfaces.

Implementation of CLARIN Type C and Type B centres does not require large coordination efforts. This can be done individually by institutions that are already hosting different linguistic resources (corpora, lexicons, etc.). Coordination would be necessary for the selection among alternative standard options, and the e-Lingua consortium could be the main driving force here.

3.3 Metadata standards for resource annotation

The main types of linguistic resources that should be incorporated into the common HLT infrastructure are corpora, machine readable dictionaries (MRD) or lexicons and linguistic databases. Therefore, the first task is to define metadata standards for the annotation of these resources.

Regarding corpus annotation standards, the three main alternatives are mentioned in the CLARIN short guide [37]: the standards developed by the International Standards Organization Technical Committee 37 Subcommittee 4 (ISO/ TC37/SC4)*, XCES (XML Corpus Encoding Standard)** and TEI P5 (Text Encoding Initiative)***. The ISO/TC37/SC4 family of standards for representing linguistic information are still far from stable. XCES, which is considered to be a de facto corpus encoding standard, is used by different national corpora (American National Corpus, IPI PAN Corpus of Polish, etc.) and is still not TEI P5 compatible, poorly documented, also rather limited in annotation levels.

* http://www.tc37sc4.org/
** http://www.xces.org/
*** http://www.tei-c.org/Guidelines/P5/
TEI P5, though being a universal standard for electronic texts, and, thus, a much more complex one, is flexible in defining different annotation levels, has well-defined semantics and rich documentation, and can be adapted to various corpus encoding needs. Similar conclusions were drawn by the maintainers of the National Corpus of Polish [38], where they selected to use TEI P5 as the encoding standard. A number of other national corpora (British, Bulgarian, Croatian national corpora, etc.) have also chosen TEI P5. TEI P5 has already been used for encoding of the largest Lithuanian corpus – Corpus of Contemporary Lithuanian Language (CCLL) [39].

Figure 1 illustrates correspondingly document level and structure level annotation of CCLL.

![Figure 1. Annotation at the document level – TEI P5 conformant header](image)

Annotation of machine-readable dictionaries or lexicons could be done via straightforward application of CLARIN recommendations. CLARIN recommends the use of ISO 24613:2008 standard “Language resource management - Lexical markup framework (LMF)” (http://www.lexicalmarkupframework.org/).

LMF defines a common model for design and use of lexical resources, that can be applied to monolingual, bilingual, and multilingual lexical resources. Same specifications can be applied for the description of small and large, simple and complex, written and spoken language lexical resources.

Annotation of semantic information normally should follow the widely accepted W3C SKOS (Simple Knowledge Organization System) standard (http://www.w3.org/2004/02/skos/). SKOS is a model, defining basic structure and content of thesaurus, classification schemes, taxonomies, controlled dictionaries and other concept schemes.

### 3.4 CLARIN Type C metadata centre – a pilot

In order to comply with CLARIN recommendations for resource access and metadata harvesting, OAI-PMH standard is the most reasonable. OAI-PMH is a specification proposed by the Open Archive Initiative (OAI) for ensuring interoperability of networked digital libraries (http://www.openarchives.org/). Repository of a data provider is OAI-compatible, if it can respond to OAI-PMH requests: GetRecord, Identify, ListIdentifiers, ListMetadataFormats, ListRecords, and ListSets. The minimum requirement is to provide at least unqualified Dublin Core (DC) metadata description for the resources. This means, that for linguistic resources, annotated with TEI P5, alternative DC annotation should be maintained.

Feasibility of OAI-PMH for TEI P5 annotated resources was tested by providing access to CCLL resources via metadata harvesting. At the same time, a pilot implementation of a CLARIN Type C compatible metadata centre was tried by applying the following scheme:

1. Resource annotation using DC metadata.
2. Implementation of OAI-PMH repository management software.
3. Resource upload, following the OAI-PMH requirements.
4. OAI-PMH request testing with OAI-PMH harvesters.

Figure 2. Response to OAI-PMH GetRecord request

Figure 2 shows the response of the pilot OAI-PMH repository to GetRecord request. Summarizing, the pilot implementation proved that CLARIN recommendations are rather flexible and can be implemented on top of existing linguistic resource formats without having to make any big changes.

4 Conclusions

Existing Lithuanian HLT tools and resources are rather disparate and the task of integrating them into a service-providing is extremely important. In developing national HLT infrastructure, interoperability, standardisation and international integration issues have to be considered. With many different standards existing, the most reasonable way of Lithuanian HLT infrastructure development would be compatibility with requirements of corresponding pan-European infrastructures, where CLARIN seems to be the most promising. A pilot experiment shows, that CLARIN recommendations of standards for resource description, resource access and metadata harvesting can be flexibly applied to already existing linguistic resource description framework.

References


A NEW APPROACH FOR RASTER IMAGES ANNOTATION

Giruta Kazakeviciute - Januskeviciene¹, Eugenijus Januskevicius²
¹Vilnius Gediminas Technical University, Department of Graphical Systems, Sauletekio al. 11, Vilnius, Lithuania, giruta@vgtu.lt
²Vilnius Gediminas Technical University, Faculty of Architecture, Pylimo g. 26/1, Vilnius, Lithuania, ejs@ar.vgtu.lt

Abstract. Growth of media data creates huge problems when searching and classifying the depicted objects. Currently search is based either on surrounding context, which is irrelevant in most cases, provided image meta data sets or low or mid-level regional peculiarities of digital images. The article is focused on a specific case of image meta data - annotations: textual elements, describing or explaining objects, depicted in the image. The paper presents novel approach of data storage in the field of digital imaging, embedding annotation data in format independent way directly into digital images. The annotation scheme presented is targeted to JPEG-2000 as upcoming image storage and transmission standard, but can be easily incorporated into other image storage formats, both lossless and lossy. The scheme allows wide taxonomy of images to be annotated: digital photos, aerial photos, scanned material, medical data sets, etc. The scheme keeps visual and structural fidelity of digital images, exploring properties of human visual system and wavelet transformation. The primary aims of the presented scheme is sharing and transfer of knowledge, like preparing material by educators, but can be extended to broader fields of application, like video, data mining or replacement for image tags and descriptions, thus creating true searchable media environment.

Keywords: annotation, data embedding, digital image, fidelity.

1 Introduction

Nowadays a large and growing amount of information is stored in various multimedia formats – as images, video, audio. A common and successful approach to organize and manage huge quantities of information is to enrich the media with meta data, tags, labels and annotations. Textual information attached is not only a way to organize information, but it enables the creation of a knowledge base or knowledge base systems for different applications, can be used for knowledge discovery and information mining from task-based imagery. It is said: “We are drowning in information but starved for knowledge ” (John Naisbitt).

Image annotation is multidisciplinary area and related with computer vision, data processing, data mining (also called knowledge discovery and data mining – KDD), artificial intelligence, machine learning and more others. It is widely acknowledged that image annotation is an open and very difficult problem in information technology science. Image annotations methods depends on images application, on image type and format, on image storage repository.

There are numerous users for which precise image annotation is very important: researchers, medics, students storing images on personal computers hard drives, ordinary users inside an organization, etc. These users require not only an accurate description of the image, like author, date, time etc., contained in meta data, but also domain specific knowledge. For this purpose the annotations should be performed in a standardized and constrained manner. Typically, in [2], [5], Error: Reference source not found, [11] content has been represented using a strict ontology to ensure that the information is consistent. This is the most common type of annotation in the semantic web. Ontology for annotations can be quite complex and time-consuming. Most of the current annotation tools provide the ontology, which is displayed in the form of a large tree. Annotation is done by scrolling over a tree and selecting an element from this tree.

In other cases, the strict constraints are less relevant, for example users that simply want to annotate their own pictures for sharing with friends [12], [13]. For personal use folksonomies – free annotation are more popular. It is a way to attach emotions and memories. It is convenient when annotation of images can happen every time an image is accessed. Naturally, not all annotations must be widely available. Some annotations can be personal, others may stay within specific boundaries (an example, in a department), and others can be made available. The most frequent application, analyzed in science literature is image retrieval [2], [11], [12], sometimes management of personal digital photo collections Error: Reference source not found and medical imagery [4], [5], [9], Error: Reference source not found, rarely education, data mining, transfer of knowledge [4] and similarly, and these tasks are analyzed on the semantic WEB environment.

The primary focus of this work – on the novel approach of data storage in the field of digital imaging, embedding annotation data in format independent way directly into digital images with the purpose sharing and transfer of knowledge not only on line but off line too.

Section “State-of-the Art” presents the existing image annotation methods, thus underlying the novelty
of our approach. After description of general requirements for image annotation, the quality evaluation metric for annotated images is proposed. In section “A new annotation strategy for task-based imagery” is reviewed the new image annotation scheme. The article closes with experimental results, conclusions and future work.

2 State-of-the-Art

2.1 Methods for images annotation

The widely used paradigm for information description in digital media is the use of supplementary data – meta data. It is usually organized in pairs of attribute and value, and to some extent is self-describing. Text-based image annotation continues to be an important fundamental problem in the computer vision, information mining, retrieval communities. With rapidly increasing collections of image data on and off the Web, image context understanding, search and retrieval is becoming actual an employment of proper annotating methods for task-based imagery. The most popular image annotating methods is presented below.

2.1.1 Dissociated annotations

In the majority of works [4], [5], [9], Error: Reference source not found, [11], [12] author or user can add information to an image using Extensible Markup Language (XML), Resource Description Framework (RDF), Web Ontology Language (OWL).

XML-based annotations play an important role in describing on line images information on the Internet. XML-based annotation information is stored independently from the image data, and combined with the image data at display time. An “annotated image” is the combination of an “annotation file” and its associated image. A single image may be referenced by many annotation files. The description in RDF form can include any meta data, defined in XML scheme.

Ontology is used as the new mean of creating and using the meta data in annotation, search and retrieval. It basically contains concepts and their relationships and rules. Adding a hierarchical structure to a collection of keywords produces a taxonomy. An ontology can solve the problem when some keywords are unclear. Ontologies are important for the Semantic Web, and a number of languages, such as OWL and RDF, exist for their formalization. Formats like DIG35 [32] or MPEG7 [33] are milestones to create and handle images descriptions in a flexible way. They are based on XML-description scheme [34] and believed to be the best paradigms to handle structured data. Here information is described and organized by using extendable, hierarchical and freely-definable structures.

Despite the good properties of XML-based approaches, annotations of such type can be lost if the image file format or image name is changed or user is off line. When the image is transmitted XML-based annotation information create additional payload.

2.1.2 Annotations embedding in to the image format

The application of meta data is closely related with the media storage format, as usually the storage format defines the options for meta data storage and retrieval. Basic implementations of this paradigm can be found in applications where the available information for the description is supposed to be fixed and non-expandable. Typical examples are the EXIF information, used in digital cameras, or Dublin Core Meta data standard [15]. This standard is used to add meta data to a wide variety of resources in a simple manner. Similar to the Dublin Core categories, but focusing only on images, are the Visual Resources Association (VRA) Core Categories [25].

Embedding information into an image file format can be done in a variety of ways. The simplest is the use of features provided by the file format, like JPEG (EXIF) [6] or PNG image file formats. Fixed fields are reserved for attribute values and might be filled with the current description, whereby quantity and kind of attributes depend directly on the file format used. Another and more flexible approach is to embed arbitrary meta data is used in modern formats like JPEG2000 [7] or in parts in the TIFF or TGA formats [8]. Here the concept of containers or boxes to store the information is used. The number of boxes is not limited and they can be filled with arbitrary content, e.g. user-defined attribute pairs.

However, all those approaches simply attach additional data to the image, creating additional payload when the image is transmitted. As the attached information is not integral part of the image data, there is no uniform support for certain attributes in every image file format, and most of the description is lost if the file format is changed.

2.2 Information embedding in to the image content

The idea is related with information hiding techniques, such as watermarking and steganography. Traditionally, steganography have been designed for security-related applications and watermarking, like
copyright protection and data authentication. Recently, researchers have attempted to broaden information hiding application scope to other oriented applications [35], [36], [1], [37].

2.2.1 Visible annotations

Visible annotations as visible watermarks is the standard part of the image data but they clutter the image data and decrease visibility of the depicted objects on the image. These type of annotations is the most suitable for large images, like panoramic, satellite and similar [14]. There has been a plurality of literature focusing on label placement on maps beginning as early as 1962 with [38] and continuing today. Many systems introduce idea hiding objects when they are below a certain minimal magnification threshold.

2.2.2 Invisible annotations

The advantages of invisible annotation approach are that the embedded data becomes an inseparable part of the media and takes almost no additional storage space.

Many algorithms applied in steganography might also be suitable for embedding of meta data, since the general focus to emphasize capacity is one of main requirements for media description. Nevertheless, these approaches might also be adapted to further increase capacity by omitting the requirement of undetectability. Many of these approaches are designed for the spatial domain as it provides more capacity and controllability for embedding and extraction of the additional information. Here, the most common approach to embed data is to substitute certain bits of the available pixel representation by bits of the additional data [39]. This technique is known as LSB-approach in steganography.

Since it capacity is inversely proportional to robustness [40], these approaches are highly vulnerable if the image is manipulated after embedding. Nowadays, image data is mostly used in compressed form to reduce storage space and transmission costs, which often includes a lossy transformation of the image content. Due to their special alignment, watermarking techniques are more robust against such manipulations than steganographic approaches [40]. Most of them have been designed for a certain transformation domain and embed the additional data (watermark) directly into the transformed image representation to be resistant against modifications imposed by the transformation itself. Dependent on the used transformation domain, e.g. Discrete Cosines Transform (DCT) [41] or Discrete Wavelet Transform (DWT) [42], or specific file formats using transformation domains, e.g. JPEG Error: Reference source not found or JPEG2000 Error: Reference source notfound, many different techniques have been proposed. Nevertheless, a general problem with all of these approaches is, the capacity is in general very low and the whole image is used for embedding to reduce the risk the embedded data can be simply removed by cropping parts of the image. To create a reasonable technique for media embedding relevant properties and approaches from steganography and watermarking approaches must be combined and enhanced.

A combined technique derived from steganography, is the LSB-approach applied to DWT coefficients. Bit plane modifications usually have the highest possible informational capacity but they are not robust to even the slightest modifications. As media description goal is to maximize capacity, we can adopt the LSB approach, keeping in mind its disadvantages. Given the general analyzes of current annotation methods, it is clear there is much room for improvement in the state-of-the-art.

2.3 General requirements for image annotation

As the focus is on embedding data directly into a host image, the requirements of approach are highly related with steganography and watermarking, but has some differences [1].

The goal of an image annotation – provide additional cognitive information to the information, already presented in the image. Typical use scenarios include:

- providing descriptive information on separate objects of the image;
- revealing image peculiarities or the image or peculiarities of the objects depicted;
- providing non-obstructive way of descriptive data presentation;

Typical usage fields include:

- professional users, like instructors, using annotations to present the material; medical practitioners using annotations to denominate peculiarities of medical images; researchers and scientists, using annotations to spread the revealed knowledge;
- non-professional users, like home uses, using annotations to reveal objects depicted.

The annotated image is as useful as the image data is useful, so image data, but not the embedded information must be protected. This means that an annotation, its length or quality can be sacrificed over the quality of the digital image.

As the format of image annotation is open, there is potential risk of annotation be removed or altered.
We do not assume annotation will ever be used to provide vital data, so any alteration or removal of annotation is allowed. To provide necessary level of security, data-level security methods, like digital signatures, may be used.

Since there is no limitation regarding the image content to be described, shape and number of image regions, and the length of each description are a priori only limited by the properties of the host image. It is obvious that the sum of all annotations cannot exceed its provided capacity. Nevertheless, this is a strong demand assuming an optimal and redundancy-free usage of the provided capacity, which is rather hard to fulfill for arbitrary application fields.

Based on this assumptions, we are able to derive basic requirements for embedding region-based image descriptions into raster images:

1. The scheme is targeted to off-line users, or users, having limited data connection capabilities. The limitation is imposed to be able to use annotated data in situations when connection an external database or network node is limited for security reasons, connection costs or lack of infrastructure.
2. All embedded data belonging to regions shown on screen must be retrievable without any additional transmission of any data.
3. The extraction should not depend on current image quality.
4. The need to transmit annotations separately must be avoided.
5. Annotations may contain any binary data.
6. Regions to be described may have any shape.
7. Number and length of embedded data is limited by the provided capacity and quality of the host image.

2.4 Quality evaluation of annotated images

A digital imaging system involves many components: acquisition and storage, image processing and analysis, compression, printing and display. Each of these components can influence the final digital image quality. Accordingly, with the rapid development of imaging systems, have been developed the metrics of images quality.

We are interested mostly in image degradation during image compression and processing (transformed image data), because it is in our scope of this work.

Because of the differences in the characteristics of image application and contents, image formats and users, existing objective natural image quality metrics does not provide satisfactory results. Standard quantitative image quality metrics, such as peak signal to noise ratio (PSNR) and mean squared error (MSE), are not directly related to human perception. For the image quality evaluation we need automated objective quality assessment methods that are guided by the human vision model in order to accurately reflect human perception.

In the most image processing environments human eyes are the ultimate receivers of images. For many years the subjective quality measurement Mean Opinion Score (MOS) has been used. However, subjective evaluation is usually too inconvenient, time-consuming and expensive in practice.

Many suggestions have been made to develop new objective image quality measurement, incorporating perceptual quality measures by considering human visual system (HVS) characteristics [16], [17], [18]. But in [16], [23] has been noticed that none of the complicated objective image quality metrics has shown any clear advantage over simple mathematical measures such as PSNR under strict testing conditions using different image distortions. Later, in [19], a consortium of experts, the video quality experts group (VQEG) [20] reported that none of the new proposed methods, tested under different configurations, were comparable to the “null mode,” a hypothetical model that predicts quality exactly. It means QA methods needed to improve further.

The perceptibility of image details depends on the sampling density of the image signal, the distance from the image plane to the observer, and the perceptual capability of the observer’s visual system [22]. The subjective evaluation of the images varies when these factors vary. A single-scale method (SSIM) is appropriate only for specific settings. Using multi-scale method (M-SSIM) is possible to incorporate image details at different resolutions.

The authors of [22] have made the quality assessment (QA) algorithms freely available to the research community [24]. This has allow other researchers to report their comparative results on widely accepted QA algorithms. In [30], [31] have presented a novel information theoretic criterion – visual information fidelity (VIF) criterion for natural scenes. VIF was developed for image and video quality measurement based on natural scene statistics (NSS). Images and videos of the three-dimensional (3-D) visual environment captured using high-quality capture devices operating in the visual spectrum are classified as natural scenes [30]. Images and videos captured from non-visual stimulus such as radar and sonar, X-rays, ultrasounds, computer generated graphics, cartoons and animations, paintings and drawings, random noise, etc. differ from natural scenes.

By combining information content weighting with multi-scale SSIM, in [26] has been defined an information content weighted SSIM measure (IW-SSIM). Authors of this research has looked at the image QA pooling problem from an information theoretic point of view. In computational vision science is widely
hypothesized that the human visual system (HVS) is an optimal information extractor [27]. To achieve this optimality, the image components that contain more information content would attract more visual attention [28]. The local information content can be quantified in units of bit. Such statistical image model is available and then the local information content measure can be employed for image QA weighting.

As we see from the recent extensive tests results with six publicly available independent image databases [29], the proposed IW-SSIM algorithm achieves the best overall performance.

3 A new annotation strategy for task-based imagery

The annotation scheme we are proposing is targeted towards communication and transmission of knowledge. We assume the possibility to annotate a depicted object is valued more than possibility to secure the annotated information. Because of this, security issues, like removing annotation or altering its contents are intentionally left aside.

The scheme (Fig. 1) proposed uses meta data to describe the spatial position of the Region of Annotation (ROA) in the image and textual or binary data to describe the contents of ROA (see Fig. 3).

The ROA can be of any geometrical shape. The only limitation set is ROA must fully be contained in the image or image part. If the same annotation is to be placed in different image parts, like tiles in JPEG 2000, every part must contain its own ROA. The coordinates and size of the shape are defined in spatial coordinate system of the initial image or image part. 3 basic shaped are defined now:

- rectangular, described by upper left corner, width and height;
- circular, described by center point and radius;
- free-shaped, described by a set of points, defining its boundary.

The data for annotation can be either textual, using Unicode as symbol set and allowing to embed any glyph, or binary, allowing any binary data, like image or sound, to be embedded into the carrier.

The scheme utilizes Laplacian pyramid decomposition, allowing compact annotation storage and tight spatial relation with the ROA (see Fig. 4). The annotation can be stored in the same spatial region as ROA or can be detached and stored elsewhere in the same image. There are two aims of annotation detaching:

- preservation of quality in ROA, needed in some image taxonomies, like medical images;
- increasing capacity of annotation, when it does not physically fit within ROA.

The proposed annotation scheme is targeted to high information density, so standard data description
schemes, like SGML or XML cannot be used. Instead of this, a hash, of “key-value” approach is used (see Fig. 6). This allows to store information in more compact form, because the total overhead of XML format for small data set can be as high as 3000% (see Fig. 2) but it tends to decreases while increasing the amount of useful information.

The annotation, containing both spatial information and annotating data, is packed into structure, suitable for embedding. The structure contains information of the annotation, graphical finder patterns and error-correcting codes. The graphical shape of the embedded structure reassembles Datamatrix - a family of 2D bar codes (see Fig 5).

![Figure 5. Various instances of embedded annotation](image)

Figure 5. Various instances of embedded annotation

![Figure 6. Data structures used for creating graphics in Fig. 2. Useful information is marked in boldface](image)

Figure 6. Data structures used for creating graphics in Fig. 2. Useful information is marked in boldface

![Figure 7. Bitplanes: original (a and b) and with embedded data (c and d)](image)

Figure 7. Bitplanes: original (a and b) and with embedded data (c and d)

The initial selection of embedding method is based on the lowest possible impact on information being annotated, low computational requirements and possibility to change the impact on the quality in the most flexible way. The embedding method chosen is a combination of LSB modifications and QIM and works in the lower bit planes of quantization indexes. Figure 7 shows bitplanes of HH subband before a) and b) and after embedding c) and d). Due to high entropy of the lower bit-planes, the embedded information is almost indistinguishable from the coefficient data.

The proposed scheme is targeted to JPEG 2000 image storage format and perfectly fits into standard JPEG 2000 image coding and decoding pipeline. Data and image preparation methods (see Fig. 1) for annotation are handled in cost-efficient way and can be re-used in any image storage format with negligible computational overload.

4 Experimental results

The proposed image annotation scheme was tested in laboratory conditions. To test annotation scheme, several publicly available image databases were used. Test images selected are from different categories, at least 3 images were used from each category. Testing goal was to measure visual impact on the image, caused by annotations. Only gray-scale images were tested, mainly because IW-SSIM index is defined for gray-scale images only.

![Figure 8. Image quality measures: original uncompressed image against the compressed image containing annotations](image)

Figure 8. Image quality measures: original uncompressed image against the compressed image containing annotations

![Figure 9. Image quality measures: compressed image against the compressed image containing annotations](image)

Figure 9. Image quality measures: compressed image against the compressed image containing annotations

Testing consists of 2 stages. The first stage consists of a single ROA, having size of 40-by-40 pixels. The payload for annotation is the first sentence from “Lorem ipsum...” containing 124 bytes. The second stage contains 5 annotations, each containing the same first sentence from “Lorem ipsum...”. The third stage contains 10 annotations, containing first 50 symbols from “Lorem ipsum...”. The results, presented in Figures 8 and 9 depicts average values of these tests.
IW-SIM, PSNR and MS-SIM values were measured. Additionally JPEG 2000 compression was introduced, varying compression ratios from 1:1 to 200:1. Maximal compression ratio taken was 200:1 for topographical images and 50:1 for medical images. Visual impact of embedded annotations was measured between the original uncompressed image and the compressed image containing annotations (set 1) as well as between compressed image and the compressed image containing annotations (set 2). The figures below presents average PSNR, IW-SSIM and MS-SSIM quality measurement for set 1 (left side) and 2 (right side) values for every type if image taxonomy tested. PSNR value is presented for reference only. The decision on image suitability for particular use should be made relying on either IW-SSIM, or MS-SSIM values. In this project we assume IW-SSIM value of 0,85 is the lowest suitable for natural and areal images, while 0,90 - for medical images in most modalities.

As presented by measurement series 2 in Figure 8, influence of moderate information embedding is almost unnoticeable for the user and image quality. The increase of image quality from compression ratio 1:1 to 1:2 or 1:4 should be explained by different DWT filters used. Average retrieval rate of embedded information in this stage is 70%. This can be improved by using more sophisticated methods of information embedding, as mentioned in “Future work” section.

5 Conclusions and future work

The proposed approach is suitable for information embedding in JPEG-2000 imaginary in small and medium quantities. The embedding process can be integrated in the process of JPEG-2000 encoding ant take almost no additional computational time. Visual impact on the compressed image is neglectable when compared with the compression artifacts.

The analysis of retrieval process reveals the information is lost due to peculiarities of JPEG-2000 encoding process, when targeted bitrate is nearly reached. To improve the quality of information retrieval, a lifting scheme, used in ROI encoding could be used.

Color images can be used for information embedding, having suitable quality measurement methods. The targeted component is “Y”, or luminosity component. The choice is based on analysis of test results, as the quality of image does not suffer noticeably while embedding the information. Moreover, the Y component exposes just a little more entropy, as it is compressed less in comparison to U and V components.

References


* Some modalities does not allow DWT-based compression, while permitting DCT-based compression.
SVM AND XBRL BASED DECISION SUPPORT SYSTEM FOR CREDIT RISK EVALUATION

Paulius Danenas, Gintautas Garsva

Vilnius University, Kaunas Faculty of Humanities, Department of Informatics, Muitines St. 8, Kaunas, Lithuania, paulius.danenas@vukhf.lt, gintautas.garsva@vukhf.lt

Abstract. This article presents a framework for an intelligent system for credit risk evaluation and a model for its implementation in cross-platform and database-independent way. This framework supports widely applied Support Vector Machines (SVM) technique as basis for intelligent evaluation and eXtensible Business Reporting Language (XBRL) standard which is widely developed, implemented and supported by governments and regulatory authorities. It also examines current SVM frameworks which can be used for new hybrid method development and the possibilities to implement them in a credit risk decision support system. A framework for a distributed credit risk evaluation system is presented together with its architecture and model for implementation.

Keywords: Support Vector Machines, SVM, XBRL, XML, artificial intelligence, machine learning, credit risk, evaluation, bankruptcy, framework, decision support system, DSS.

1 Introduction

Credit risk evaluation currently is a hot topic because of large number of companies which are going bankrupt, thus the price of mistake by approving credit request or rejecting it is high. This is why it is very important for the creditor to choose the correct principle or model for evaluation. Tools such as artificial intelligence and soft computing offer a possibility to create more sophisticated and precise models which can be implemented in such systems. Support Vector Machines (SVM) is one of these methods widely applied as an effective solution to many various pattern recognition, classification, regression and forecasting problems, including financial forecasting and credit risk evaluation. SVM technique has proven itself as an effective solution in credit risk field with results comparable to or better than most of other machine learning techniques [1]. SVM method has been combined with almost all popular natural computing techniques – fuzzy logic [2][3], Bayesian inference[4], genetic algorithms[5], rough sets [6], ant colony optimization [6], particle swarm optimization [7]. A research which combines SVM with widely used discriminant techniques to generate classification rules has also recently been made [8]. These investigations together with many others proved that SVM-based methods often outperformed similar techniques, such as Neural Networks, and can be deployed as a solution in credit risk field.

Expert systems and decision support systems (DSS) implementing these techniques are also widely used to help to solve this problem. They are widely researched and developed. There are many researches targeted to expert systems and their application in credit risk field; from rules-based using inference [9] to model-based with particular models [10][11], or are related more to specific tasks in credit risk field rather than for credit risk evaluation, e.g., credit card assessment [12]. Some researchers implement multi-agent based solutions as a network of problem solvers that perform together to solve problems [13], other focus on decision support system architecture with application of modern machine learning techniques such as SOM [14] or SVM [15]. Earlier developments also apply multiple criteria decision aid [16], as in most of similar researches financial statements and various financial ratios are used as main source of information. However, there seems to be a lack of research for integration of modern financial standards which offer new possibilities for financial and intelligent evaluation. The aim of this research is to propose a structure for decision support system for credit risk evaluation which implements advanced technologies and techniques such as Support Vector Machines (SVM) and eXtensive Business Reporting Language (XBRL).

2 SVM method and its implementations

SVM was created in the 7th decade by Vapnik in the labs of AT&T Bell Company. SVM and some other machine learning methods are a part of statistical learning theory, which describes the characteristics of self-learning machines. This theory gives theoretical foundations not only for SVM, but also for other important machine learning methods, such as neural networks. Vapnik defines SVM as learning machines that can do binary classification (pattern recognition) and real valued function approximation (regression evaluation). SVM implicitly maps m-dimension input data space to n-dimension possibility space where a linear classifier is created. The special ability is that at the same time it minimizes empiric classification error and maximizes geometric margin. This is the reason why this method is called maximum margin classifier.
The main task of binary classifier is to evaluate function $f: X \rightarrow \{±1\}$ which maps input and output data. According to Vapnik (1995), SVM can be formulated as if empirical data $(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n) \in \mathcal{X} \times \{±1\}$ is given with $x_i \in R^d, y_i \in \{-1,1\}$, the task is to find a decision function $f_{w,b}$ with the property $f_{w,b}(x_i) = y_i$, $i = 1..n$. The similarity measures for $\mathcal{X}$ and $y$ can be formally described as a function $k$, commonly referred as kernel function: $k: \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$, which returns a numerical characterization of the similarity between $x$ and $x'$. It is used to transform input data space into another vector space (usually in much larger number of dimensions) in such a way that transformed data space can be separated in a linear way. The most popular and applied kernel functions are linear, polynomial, radial basis function (RBF) and sigmoid (as in other kernel methods, such as Neural Networks). SVM offers such advantages, according to Scholkopf [17]: similarity measures based on dot product in space $F$, interpretation based on analytical geometry, a possibility to select nonlinear mapping $\Phi$ for more suitable data representation.

There are many SVM algorithms and software for further classifier development available. Specialized knowledge in data mining and machine learning, as well as knowledge in particular field is required for creating hybrid models. A comparison of the capabilities of these frameworks is represented in Table 1. The comparison includes LibSVM[18], LS-SVM[19], also Lagrangian SVM[20], ASVM[21], SSVM[22], LPSVM[23] and Proximal SVM[24] by Mangasarian et al., SVMLight[25], BSVM[26], UniverSVM[27], SVM&KM Toolbox[28], SimpleMKL[29], mySVM[29], TinySVM[30], Core Vector Machines[32], PSVM[33], GPDT[34] and LIBLINEAR[35].

Table 1. Comparison of SVM implementations

<table>
<thead>
<tr>
<th>Problems solved</th>
<th>LibSVM</th>
<th>BSVM</th>
<th>UniverSVM</th>
<th>mySVM</th>
<th>TinySVM</th>
<th>SVM &amp; KM Toolbox</th>
<th>SimpleMKL</th>
<th>SVMlight</th>
<th>CVM</th>
<th>PSVM</th>
<th>GPDT</th>
<th>LSSVM</th>
<th>Lagrangian SVM</th>
<th>ASVM</th>
<th>SSVM</th>
<th>LPSVM</th>
<th>Proximal SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Regression</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Ranking</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Clustering</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Feature selection</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
</tbody>
</table>

| Number of classes | One-class | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Number of classes | Two-class | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Number of classes | Multi-class | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |

| Programming language | JAVA | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Programming language | MATLAB | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Programming language | C/C++ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Programming language | Python | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Programming language | Perl | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |

| Kernel function | Linear | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Kernel function | Polynomial | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Kernel function | Sigmoid | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Kernel function | RBF | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Kernel function | ANOVA | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Kernel function | Neural | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Kernel function | Laplas | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Kernel function | User-defined | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Kernel function | Point | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Kernel function | Sum | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Kernel function | Inverted distance | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |

As SVM-based methods are widely applied in research, data mining and data analysis, SVM implementations are almost in every professional or scientific package for statistics or data mining (e.g., Weka[36] and RapidMiner[37], which have interfaces with popular SVM packages LibSVM, LIBLINEAR). However, currently there are no professional or scientific toolkits built exclusively for SVM method which would include all or most of the SVM frameworks mentioned in Table 1; the closest to the solution of this problem is Shogun[38] framework, but it is seems to be complicated to compile and use it on non-Unix OS. This
is important as different SVM classifiers might produce different results thus it might be useful develop classifier using several SVM methods.

3 XBRL standard

XBRL (eXtensible Business Reporting Language) is an XML-based standard for financial reporting currently is becoming one of the most important frameworks in financial and banking sectors. This framework defines a widely used uniform and system-independent standard and utilizes the possibilities of XSL/XSLT/XSL-FO technologies, which can be applied to transform XML documents to RTF or PDF formats as well as other XML documents. It is applied as a standard in such countries as USA (Securities And Exchange Commission adopted it in 2005), UK (Financial Services Authority started using it in 2005). CEBS (The Committee of European Banking Supervisors) has also applied it as standard for 9000 credit institutions and investment firms to 25 supervisors/ regulators[38]. This standard is also the basis of new SEC’s financial EDGAR database implementation. Great Britain will also apply XBRL based framework as a financial standard in 2011 April as all company tax returns will be bound to be filed online using the new iXBRL Standard[40].

XBRL is defined by two primary concepts: taxonomy and instance. Taxonomy defines all financial concepts that are used by a particular entity, as well as their inner relationships and internal or external resources; instance can be defined as the list of facts which has the structure defined in taxonomy. Semantic meaning is expressed by fully leveraging XLink to connect XML instance documents with any number of related XML fragments at a granular level using extended links. XBRL taxonomy consists of taxonomy and linkbase components; additional module specifications are also defined, starting from XBRL 2.1 version. These components are listed in Table 2.

Table 2. XBRL structure

<table>
<thead>
<tr>
<th>Component</th>
<th>Concepts stored (purpose)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core</td>
<td></td>
</tr>
<tr>
<td>Taxonomy</td>
<td>Unstructured list of elements and references to linkbase files</td>
</tr>
<tr>
<td>Linkbase</td>
<td></td>
</tr>
<tr>
<td>Presentation</td>
<td>Relationships between elements in order to properly organize the taxonomy content (representation of the hierarchical relationships in particular business data)</td>
</tr>
<tr>
<td>Calculation</td>
<td>Definitions of basic validation rules for a particular taxonomy. Might be used to sort all monetary elements hierarchically (upper level entity is the result of operation with lower level elements)</td>
</tr>
<tr>
<td>Definition</td>
<td>Different kinds of relations between elements (distinguishing “general-special”, similarity “essence-alias”, requirement “requires-element”, tuple similarity “similar-tuples”)</td>
</tr>
<tr>
<td>Label</td>
<td>Labels for different languages</td>
</tr>
<tr>
<td>Reference</td>
<td>Pointers to source documents which describe the concepts defined in taxonomy</td>
</tr>
<tr>
<td>Additional modules (specifications)</td>
<td></td>
</tr>
<tr>
<td>XBRL Dimensions 1.0 [41]</td>
<td>Definition of additional structured contextual information for business facts in a manner similar to „dimension” concept in OLAP analysis. The base XBRL Dimensions specification defines three dimensions: reporting period, reporting entity (e.g.; a company or its division), and a loosely-defined reporting scenario, originally intended to distinguish between actual vs. projected facts. Taxonomies using XBRL Dimensions can define new dimensions as well as relate other taxonomy metadata (labels, presentation information, etc.) to them.</td>
</tr>
<tr>
<td>XBRL Formula 1.0 [42]</td>
<td>Definition of formulaic expressions using XPath 2.0 to validate XBRL instance information or to derive new XBRL facts (e.g. calculate secondary financial ratios)</td>
</tr>
<tr>
<td>XBRL Rendering [43]</td>
<td>Provision of documents which can be viewed in a browser while making use of XBRL tags and leveraging label and presentation linkbases to provide more comprehensive report definition, e.g., to specify exact relative location of particular concept</td>
</tr>
<tr>
<td>XBRL Versioning [44] (still in development)</td>
<td>Documentation of the differences between two versions of the same taxonomy (i.e., further taxonomy development, extension with new languages or business rules, errors found, laws and regulations that have changed, change of concepts or their removal and etc.). It should also provide ability to ensure compatibility for software vendors</td>
</tr>
</tbody>
</table>

SEC defines a particular taxonomy for Nationally Recognized Statistical Rating Organizations (NRSRO). It consists of 11 files and differs in absence of calculation linkbase, but additionally has type and role declaration schemas. The following aspects are defined in the latest version (currently Ratings 2009 Taxonomy): SEC category type (Corporate, Insurance, Finance, AssetBacked and Government), SEC subcategory type (Sovereign, US Public Finance, International Public Finance), Ratings item type (Program, Instrument, Shelf, Other), Instrument identifier scheme type (ISIN, SEDOL, VALOR, WKN, SICC, NRSRO, Other) and Issuer identifier scheme type (CISIP, DUNS, BIC, SICC, NRSRO, Other)[45].

Several public taxonomies are developed as authority standards; mostly applied are US GAAP used in USA as a standard for SEC reporting. Other well known taxonomies are COREP – Common Reporting taxonomy created by The Committee of European Banking Supervisors for solvency ratio reporting in Europe.
under future EU capital requirements regime and FINREP – Financial Reporting taxonomy created by The Committee of European Banking Supervisors, using Belgium Banking Taxonomy as basis. COREP is part of Basel II standard [38].

4 Proposed framework for credit evaluation decision support system

The human-system interaction is modeled by Use-Case model (Figure 1) which defines the main roles that use the particular functions. Two main roles are defined, Analyst and Risk Manager; however, it may depend on the structure or specifics of credit organization. The functions that they perform are targeted primarily at model creation and evaluation process as well as applying the results in practical activity and data management. The system is modeled as distributed system thus a modeling service subsystem is defined which represent the main classifier operations: model training (creation), testing and prediction. Model creation task is defined as a complex process which includes such subtasks as parameter selection (using stochastic or evolutionary techniques or selecting them manually), data preprocessing, training, testing and saving model to model repository. Data analysis can be viewed as a generic process that can be extended to statistical, financial or visual analysis. Yet analysis in credit risk is a complex task which needs all three ways of evaluation so this task can also be viewed as an aggregation of these three.

The structure of intelligent decision support system in presented in Figure 2. It has main components which were defined in earlier works (e.g., [10], [11], [14], [15]) – model repository, data store and user interface. However, additional possibilities are defined for data retrieval and processing as modern XML-based standards allow creating automatic data import from various sources.

The system’s structure is defined as consisting of three main layers – SVM based machine learning layer (further referred as SVM-ML layer) which purpose is to define and implement all the machine learning techniques and algorithms necessary for evaluation as well as other data mining tasks which need to be solved in machine learning process such as information processing, representation; data layer defining data that is available for modeling and stored in data storage facility; and credit risk evaluation layer (further referred as CRE layer) that implements whole analysis, modeling, forecasting and evaluation logic, as well as data visualization. The separation of these aspects gives a possibility to use machine learning techniques implemented in this system to solve other problems by implementing only the logic specific to these problems. The main aspects of this system are also defined as particular layers:

Data source interaction layer – it is defined in both SVM-ML and CRE layers. SVM-ML layer interaction sublayer includes database interaction layer with object persistence frameworks (such as Hibernate) and database connection frameworks (drivers), as well as data standards commonly used in machine learning software (such as Weka ARFF, Comma Separated Values (CSV) formats or interoperable Predictive Model Markup Language (PMML) standard). It also defines the interfaces for intelligent information retrieval (not necessarily financial) using Web Services or intelligent agents. CRE layer additionally defines (or can be thought as an extension of previously described layer) financial standards and data sources specifically for finance or credit risk related tasks. It also has a mapping package that contains the mappings between XBRL (or other standards) and data stored in Data Layer. The XBRL instances might be entered manually or automatically,
by using, e.g., RSS feeds which contain links to these instances thus automating data retrieval process. RSS is also proposed by SEC as an option[45].

Figure 2. The structure of intelligent credit risk evaluation DSS

**Information Processing layer** – also defined in both SVM-ML and CRE layers. It implements main tasks that are solved during the whole intelligent model creation process before training using one of the algorithms. It defines such standard data mining aspects as information retrieval, data extraction and cleansing, data transformation (e.g., using Principal Component Analysis, Independent Component Analysis, Factor Analysis, etc.), normalization/standardization, data imputation. The same layer defined in CRE layer implements specific tasks, e.g., specific transformations, data transformation to absolute or percentage changes between particular ratios during particular period and etc.

**Data Layer** – defines all the data that is stored in data store (multidimensional data warehouse, database or other source). The system described here uses company data, financial data (data extracted from financial reports), company management and personnel data, historical records, market data, also macroeconomic and statistical data for macroeconomic environment evaluation (this type of data for analysis is also defined in earlier works, e.g., [10], [11] or [14]). It also contains metadata, such as reference or multilanguage data, as well as financial ratings and historical information. As it was mentioned before, ratings data can also be retrieved from XBRL documents with ratings data (e.g., instances of SEC RATINGS taxonomy). The last component is a model repository which contains all intelligent, statistical or other models (including SVM-based or hybrid models), as well as their execution log, evaluation results and their metadata.

Source: created by authors, using [11][14]
**Representation layer** - this module includes all methods and operations which are used for representation and visualization of results. It is more generic in case of SVM-ML (defines standard representations of training, testing and prediction results as well as their visualizations). CRE layer defines more sophisticated modules such as OLAP analysis together with representation of financial analysis, simulation/modeling and forecasting as well as data management functionality;

**Financial Analysis, Modeling and Forecasting** modules are defined particularly for risk evaluation layer as they present analytics, simulations and forecasting of particular domain. Here the list of methods is not complete; it can be extended by using various modern fields and methods of computational finance.

### 4.1 A possible implementation scenario

Figure 3 represents an UML implementation diagram for DSS based on proposed framework. It describes all server nodes, execution environments as well as possible technologies that can be used to implement the described system. As Table 1 shows, SVM frameworks are implemented in different languages. This makes a difficult task to combine them in hybrid algorithms or use together in single system. Thus a framework for interoperability, such as CORBA, COM or Web Services, has to be implemented in this system to ensure that as many SVM implementations might be used as possible; as it was mentioned in related work, different SVM algorithms often show different results. Web Services was selected to implement the SVM classifiers as services in this case, as it ensures maximum compatibility and is easier to implement than CORBA because known CORBA open source implementations seem difficult to apply (i.e., they do not have good reference material covering all aspects or are difficult to implement in cross-platform manner), and COM is not cross-platform. Three operations most commonly used in classification tasks are defined in WSDL document: Training, Testing and Prediction. Yet, it can be extended with other operations, as some SVM implementations also offer additional functionality, e.g., ranking (SVM-Light) or outlier detection using one-class SVM.

JAVA was chosen as an implementation language for the whole system as it offers many possibilities and frameworks needed to implement functionality described here; e.g., WEKA and RapidMiner can be a good choice to implement Machine Learning layer functionality as they contain implementations of mostly referenced SVM algorithms together with many others. There are several known open source cross-platform JAVA XBRL implementation frameworks (xBreeze Open Source Edition provided by UBMatrix, Batavia XBRL Java Library, XBRLAPI.org) which would allow to develop XBRL Import Component. It is also useful to store retrieved XBRL data in XML native database which also provides search and querying facilities and can be used as storage facility of original XBRL instances. A similar component might also be implemented for other similar...
XML-based formats, e.g., MDDL* and SDMX** standards might be utilized for macroeconomic data import, RIXML*** - for research related data import (depending on the subset defined in this structure). JAVA also offers good possibilities for enterprise-level development and integration with other systems (e.g., using ESB and JMS for messaging-oriented integration development) as well as web interface implementation and development of Web services. If SVM classifier cannot be implemented in JAVA then C++ can be a good option as Table 1 shows that most of them are written in C/C++ or MATLAB (using its mex compiler). These aspects are defined in proposed system framework. Such features of implementation proposed in Figure 3 can be emphasized:

- Cross-platform – XBRL and other XML-based standards do not depend on any system platform; JAVA can be deployed on any Windows/*nix/BSD platform and C++ also has popular cross-platform frameworks for development (e.g., Qt framework by Nokia);
- Database-independent – application of Hibernate object persistence technology allows implementing the system almost independent on DBMS as SQL queries in code are expressed using Hibernate’s own HQL technology and then translated to corresponding SQL dialect.

4.2 Comparison with similar systems and possible future extensions

Tsaih et al. proposed an N-tier architecture with internal credit scoring model transformation into XML document. It consists of thin client layer (representing GUI in Web browser), middle tiers include the web server, Management Application Server (MAS), which provides interfaces to manage TDV, Database and XML repository and perform other management tasks, Loan Processing Subsystem (LPS) with Case Processing Application Server (CPAS) and Evaluation Module (EM) with XML parser and model engine sub-modules, and Model Installing Subsystem (MIS) which consists Model Defining Application Server (MDAS) and Model Recording Module (MRM) [47]. The structure of this prototype is quite similar to proposed system’s structure, although it is targeted at real world transaction processing whereas our model is more suitable for credit risk modelling and analysis. However, its extension with industrial modules for loan processing and banking operations might be considered in the future. PNML standard might be considered as better approach for model storage, transfer and manipulation, as it offers a standardized approach and is maintained and supported by many vendors and organizations [48]. Kotsiantis et. al. developed a distributed ontology-based credit evaluation system with an application of C4.5 algorithm for scoring and intelligent search and reasoning possibilities[46]; although they referred to XBRL as one of the options which would enable analytical possibilities offered by Semantic Web technologies they chose their own developed ontology to represent financial statements. Both of proposed system prototype solutions were engineered using JAVA technologies which prove to be a good choice for implementation of such system. This framework can be extended as XBRL standard together with modern Semantic Web technologies offers many possibilities, e.g., automation of information retrieval and model updating in real time thus making evaluation even more precise. Another important extension would be business rules integration (they are defined as an artifact in Data Layer). The rules defined in XBRL Calculation Linkbase might be leveraged to ensure the integrity and validity of data as well as define additional secondary financial ratios.

5 Conclusions

Credit risk evaluation is a sophisticated task which includes many aspects, views and approaches which should be used by analyst to properly assess customer’s abilities and probability of default. Many methods are applied in this field and intelligent hybrid models based on SVM have proven themselves to be more precise and effective. A proper structure of system implementation needs be selected which would allow to implement, test and use intelligent models in credit organizations. This article proposes a structure for such system with such properties as cross-platform and database-independent. This framework is also integrated with XBRL standard which is implemented by such authorities as Securities Exchange Commission (SEC) and European Union, and is part of Basel II requirements. This standard also offers new possibilities for further research, thus, this framework might be extended with new abilities in the future.

References


* Market Data Definition Language(MDDL) (http://www.mddl.org)
** Statistical Data and Metadata Exchange (SDMX) (http://www.sdmx.org)
*** Research Information Markup Language (RIXML) (http://www.rixml.org)


[27] UniverSVM[interactive, 2010.11.30], access via the Internet http://www.kyb.mpg.de/bs/people/fabie/universvm.html


