INFORMATION SYSTEMS IN MANAGEMENT

Systemy informatyczne w zarządzaniu

Vol. 4 2015 Quarterly No. 1
Information Systems in Management
Primary version of the journal is an electronic version.

Editor
Department of Informatics, Warsaw University of Life Sciences – SGGW

Editorial Committee
Dr hab. inż. Arkadiusz Orłowski – Editor-in-Chief
Dr Piotr Łukasiewicz – Executive Editor
Dr inż. Tomasz Żąbkowski – Scientific Editor
Prof. nadzw. Kris Gaj – Linguistic Editor
Dr hab. Wiesław Szczesny – Statistical Editor

Editorial Council
Prof. dr hab. inż. Ryszard Budziński – Uniwersytet Szczeciński
Prof. dr hab. Witold Chmielarz – Uniwersytet Warszawski
Dr hab. inż. Leszek Chmielewski – Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
Prof. Jeretta Horn Nord – Oklahoma State University, USA
Dr inż. Waldemar Karsowski – Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
Prof. Yuiry Kondratenko – Black Sea State University, Ukraina
Prof. Alex Koohang – Middle Georgia State College, USA
Prof. Vassilis Kostoglou – Alexander Technological Educational Institute of Thessaloniki, Greece
Prof. dr hab. Marian Niedźwiedziński – Uniwersytet Łódzki
Dr hab. inż. Arkadiusz Orłowski – Szkoła Główna Gospodarstwa Wiejskiego w Warszawie – Chairman
Dr hab. inż. Joanna Paliszkiewicz – Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
Dr hab. inż. Jerzy Pejaś – Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
Prof. Kongkiti Phusavat – Kasetsart University Bangkok, Thailand
Prof. Josu Takala – University of Vaasa, Finland
Dr hab. inż. Ludwik Wicki – Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
Prof. dr hab. inż. Antoni Wiliński – Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

Address of the Editor
Faculty of Applied Informatics and Mathematics, WULS – SGGW
ul. Nowoursynowska 166, 02-787 Warszawa, Poland
e-mail: isim@sggw.pl, www.ism.wzim.sggw.pl

ISSN: 2084-5537

Wydawnictwo SGGW
ul. Nowoursynowska 166, 02-787 Warszawa, Poland
e-mail: wydawnictwo@sggw.pl, www.wydawnictwosggw.pl

Table of contents

Michał Bleja
ADAPTING SBA OPTIMIZATION METHODS DEVOTED TO QUERIES HAVING SUBQUERIES TYPED BY ENUMERATIONS FOR XQUERY EXPRESSIONS ................................................................. 3

Krzysztof Cabaj
VISUALIZATION AS SUPPORT FOR WEB HONEYPOT DATA ANALYSIS .................................................................................................................. 14

Renata Gabryelczyk, Arkadiusz Jurczuk
THE DIAGNOSIS OF INFORMATION POTENTIAL OF SELECTED BUSINESS PROCESS MODELLING NOTATIONS ...................... 26

Viera Gafrikova, Wiesław Szczesny, Zdzisław Odrzygożdź
ONLINE PERSONAL FINANCE MANAGEMENT APPLICATIONS ................ 39

Wioletta Grzenda
THE ADVANTAGES OF BAYESIAN METHODS OVER CLASSICAL METHODS IN THE CONTEXT OF CREDIBLE INTERVALS .......................... 53

Małgorzata Grzywińska-Rapca
EQUIPMENT LEVEL OF POLISH HOUSEHOLDS IN INFORMATION TECHNOLOGY ............................................................................................... 64

Monika Jaranowska
FUNCTIONAL ANALYSIS OF INFORMATION SYSTEMS SUPPORTING THE PROCESS OF FACILITY MANAGEMENT ........................................ 73
ADAPTING SBA OPTIMIZATION METHODS DEVOTED TO QUERIES HAVING SUBQUERIES TYPED BY ENUMERATIONS FOR XQUERY EXPRESSIONS

MICHAŁ BLEJA

Faculty of Mathematics and Computer Science, University of Łódź

The paper presents the concepts related to the design of query optimization methods for nested queries. The proposed methods are dedicated for queries having so called weakly dependent subqueries. A subquery is weakly dependent from its direct operator if it depends only on expressions typed by enumerations. We have successfully developed the weakly dependent subqueries method in the context of object-oriented database management systems based on Stack-Based Approach. Systems storing XML data which are queried using XQuery can be considered simplified object-oriented databases. For this reason we try to adopt SBQL query optimization methods to XQuery expressions.

Keywords: query optimization, XQuery, enumerations, weakly dependent subquery, Stack-Based Approach, SBQL

1. Introduction

A query optimization [1, 2] is aimed at radical reducing of the query processing time. It must take into consideration both the time needed for the optimization and the processing time after optimizing. There are many various optimization methods in contemporary database management systems. Some of them assume using redundant data structures called indices [3]. Other techniques cope with caching query results [4] and different strategies of physical data organization. The major group of methods concern query rewriting [5, 6, 7, 8].
Rewriting consists in translating an original query into a semantically equivalent form ensuring better performance.

The ODRA (Object Database for Rapid Application Development) [9, 10, 11] prototype is the main component of the European project EGov Bus [12]. It is equipped with a very powerful query optimizer. In particular, the optimizer deals with optimization methods based on query rewriting. Among them we can mention the method of weakly dependent subqueries [5, 6], pushing selection before structure operators [7], the method of independent subqueries [8]. These techniques have been developed in the context of the Stack-Based Approach [13, 14] and its query/programming language SBQL. SBA constitutes a uniform conceptual platform for object-oriented databases which uniformly covers two aspects: querying and programming. It allows to construct optimization methods in their full generality. The rewriting rules presented in [5, 6, 7, 8, 14] deal with any data model (assuming that its semantics would expressed in terms of SBA). In particular, they hold for the relational model, the XML model [15] and any version of object-oriented model. The rules cope also with any operators and make no assumptions concerning the complexity of subqueries of a given query.

The XQuery [16, 17] semantics can be defined using the SBA concepts. It requires introducing three data structures which are essential for the precise semantic description: an object store, an environment stack, and a query result stack. Besides a special phase called static analysis [19, 20] is required to equip XQuery expressions with the information essential for detecting subqueries typed by enumerations. At this moment we translate basic XQuery expressions (so called FLOWR expressions) into their SBQL equivalents. SBQL queries are rewritten by the ODRA optimizer and then converted into XQuery ones. We plan to define the XQuery semantics using SBA concepts in the future.

The research presented in this paper concerns concepts related to optimization of nested queries. We deal with a special class of subqueries of a given query referred to as weakly dependent subqueries [5, 6]. The dependency is considered in the context of query operators such as selection, quantifiers, joins, etc. A subquery is weakly dependent from its direct operator if the dependency concerns only an expression which is typed by the enumeration. The number of evaluations of such the subquery can be limited to the number of enumerators occurring in the enumeration on which it depends. For instance consider the query which returns each employee earning above the average salary calculated for all employees having his/her gender. Without optimizing the subquery calculating the average salary for genders would be evaluated hundreds or thousands of times, while it could be evaluated only 2 times (once assuming $gender = $"male" and next one assuming $gender = $"female"). This subquery is a classical example of a weakly dependent subquery.
2. The Stack-Based-Approach

The Stack-Based Approach [13, 14] presents the right theory for object-oriented databases and their query/programming languages. We present these concepts of the Stack-Based Approach which are essential to the development of optimization methods for queries having weakly dependent subqueries. SBA involves the following concepts [13, 14]:

- naming-scoping-binding - each name in a query/program is bound to a suitable run-time entity depending on the scope for it.
- environment stack - it is responsible for binding names, procedure/method calls, scope control.
- total internal identification - each entity must have a unique internal identifier.
- object relativity - objects are treated uniformly and have the same formal properties regardless of the hierarchy level at which they occur.

SBA introduces a family of object store models M0, M1, M2, and M3 [14]. The simplest is M0 which deals with relational and XML-oriented data structures. In the M0 model each object is a triple consisting of an internal identifier, an external name and a value. The M1 model extends M0 with classes and static inheritance. Classes are understood as objects which store invariants (e.g. methods) of their instances. M2 extends M1 by the concept of dynamic object role. M3 augments M2 with the encapsulation mechanism.

SBQL is described in detail in [13, 14]. The syntax of SBQL is as follows:

- A single name or a single literal is an atomic query (e.g. emp, dept, salary, "Smith", 3000).
- If \( q \) is a query and \( \theta \) (e.g. sum, avg, -) is a unary operator then \( \theta(q) \) is a query.
- If \( q_1 \) and \( q_2 \) are queries and \( \theta \) is a binary operator (e.g. where, =, +, quantifier) then \( q_1 \theta q_2 \) is a query.

XML (Extensible Markup Language) [15] is a flexible text format applied to store and exchange data. Among them we can mention Oracle XML DB, eXist-db, Apache Xindice. XQuery [16, 17] is a query language for addressing XML data. It navigates through XML documents using XPath [18] expressions. Queries in XQuery are often formulated using so called FLOWR [16, 17] (for, let, order by, where, return) expressions. To present XQuery examples we assume XML documents which correspond to the schema presented in Figure 1. For instance, the following query returns employees earning more than 3000 (The \texttt{doc} function is applied to open XML documents):

\begin{verbatim}
for $emp in doc("company.xml")//emp
where $emp/salary>3000
return $emp
\end{verbatim}

(1)
Our optimization methods are entirely performed before a query is executed. It requires a special phase called static analysis [14, 19, 20] which simulates run-time actions during compilation-time. It uses an abstract syntax tree (AST) of a given query to perform static type checking. The static analysis acts on three data structures: a metabase, a static environment stack SENVS, and a static query result stack SQRES.

```xml
<xs:schema
<xs:simpleType name="enum_gender">
   <xs:restriction base="xs:string">
      <xs:enumeration value="male"/>
      <xs:enumeration value="female"/>
   </xs:restriction>
</xs:simpleType>

<xs:simpleType name="enum_education">
   <xs:restriction base="xs:string">
      <xs:enumeration value="vocational"/>
      <xs:enumeration value="secondary"/>
      <xs:enumeration value="higher"/>
   </xs:restriction>
</xs:simpleType>

<xs:element name="emp">
   <xs:complexType>
      <xs:sequence>
         <xs:element name="fname" type="xs:string"/>
         <xs:element name="lname" type="xs:string"/>
         <xs:element name="salary" type="xs:float"/>
         <xs:element name="gender" type="enum_gender"/>
         <xs:element name="education" type="enum_education"/>
         <xs:element name="overtime" type="xs:positiveInteger"/>
      </xs:sequence>
   </xs:complexType>
</xs:element>

<xs:element name="depts">
   <xs:complexType>
      <xs:sequence>
         <xs:element name="dept" maxOccurs="50">
            <xs:complexType>
               <xs:sequence>
                  <xs:element name="name" type="xs:string"/>
                  <xs:element name="employs">
                     <xs:complexType>
                        <xs:sequence>
                           <xs:element ref="emp" maxOccurs="unbounded"/>
                        </xs:sequence>
                     </xs:complexType>
                  </xs:element>
               </xs:sequence>
            </xs:complexType>
         </xs:element>
      </xs:sequence>
   </xs:complexType>
</xs:element>
</xs:schema>
```

Figure 1. Sample XML schema
The metabase models statically the object store. It is generated from an XML schema. SENVS reflects binding operations performed on the run-time environment stack. SQRES stores signatures of the intermediate and final query results. The static analysis causes among others that [14]:

- Each name in a query is equipped with the order number of a stack section which is relevant for binding it.
- Each non-algebraic operator is assigned to the environment stack section(s) which it opens.

In SBA detecting weakly dependent subqueries is based on analyzing scoping and binding rules for names occurring in queries. A subquery is weakly dependent from its direct non-algebraic operator if it involves a name which can be statically bound to an enumerated type in the scope opened by this operator [5, 6]. Other names should be bound outside that scope. A subquery is called independent if none of its names is bound in the stack section opened by its direct operator [8, 14].

3. Optimization of queries involving weakly dependent subqueries

The approach involves the following steps:

- Translating an XQuery expression \( q_1 \) into its SBQL equivalent \( q_2 \).
- Performing static analysis of \( q_2 \) and detecting weakly dependent subqueries.
- Transforming the query \( q_2 \) into a semantically form \( q_3 \) according to the rewriting rule dedicated to weakly dependent subqueries [5, 6].
- Applying the independent subquery method [8, 14] to \( q_3 \). Let \( q_4 \) denote the result of applying this method to \( q_3 \).
- Converting \( q_4 \) into its XQuery counterpart \( q_5 \).
- Running \( q_5 \) against the underlying XML data store.

The following example presents the general idea of the above approach. The query gets each employee who has the overtime hours greater than the average overtime hours calculated for all employees having his/her education.

\[
\text{for } \text{emp in doc("company.xml")//emp}
\text{where } \text{emp/overtime} > \text{avg(doc("company.xml")//emp[education=$emp/education]/overtime)}
\text{return } \text{emp}
\]

(2)

Consider the following subquery (3) of (2):

\[
\text{avg(doc("company.xml")//emp[education=$emp/education]/overtime)}
\]

(3)

Without optimizing the subquery (3) will be evaluated once for each employee (it can be thousands of times). However it is clear that (3) can be processed only three times. The expression $emp/education$ in (3) can take only three values: occupational, secondary, and higher. We call the subquery (3) weakly
dependent because it depends from its parent query only on the expression $emp/education$ typed by the enumeration enum_education {"occupational", "secondary", "higher"}. How such a query can be detected and how a general rewriting rule should look like?

After translating the query (2) into its SBQL counterpart it takes the form (4).

For the query below we determine binding levels for its names and the number of scopes opened by the non-algebraic operators.

\[
\text{emp as } \text{$emp$ where } \text{$emp$ . overtime} > 1,
\text{avg(} (\text{emp where education=$emp$ . education).overtime})
\]

(4)

The operator as names each element in a bag or sequence returned by a query. The group as operator names the whole query result. For instance, if a query $q$ returns a collection \( \text{bag}\{e1, e2, e3, ... \} \), then a query $q$ as aux returns the collection of binders \( \text{bag}\{\text{aux}(e1), \text{aux}(e2), \text{aux}(e3), ... \} \). A query $q$ group as aux returns a single binder \( \text{aux}(\text{bag}\{e1, e2, e3, ... \}) \). The following subquery of (4)

\[
\text{avg(} (\text{emp where education=$emp$ . education).overtime})
\]

(5)

is weakly dependent from the first where operator because it contains only the name education (in the expression $emp.education$) which is statically bound to the enumeration enum_education in the scope opened by this operator. Other names in the subquery (5) are not bound in the second stack section. Denote (5) by $\text{wds}(\text{emp.education})$. Then, the original query (4) emp as $emp$ where $emp$.overtime > $\text{wds}(\text{emp.education})$ can be transformed to (6):

\[
\text{emp as } \text{$emp$ where if } \text{$emp$.education = "vocational" then } \text{$emp$.overtime} > \text{wds("vocational")}
\]

(6)

After unfolding (6) we retrieve the query (7):

\[
\text{emp as } \text{$emp$ where if } \text{$emp$.education = "vocational" then } \text{$emp$.overtime} > 1,
\text{avg(} (\text{emp where education="vocational").overtime})
\]

(7)
The general idea of the method is to construct a suitable conditional statement (if-then-else). The conditions of the statement are based on the enumerators occurring in a given enumerated type. Is the above transformation advantageous for performance? The previously considered weakly dependent subquery (5) has been replaced by the three independent subqueries: \( wds(\text{"vocational"}) \), \( wds(\text{"secondary"}) \), and \( wds(\text{"higher"}) \). Indeed, all three subqueries of (7) are independent from the first where operator. None of their names is bound in the second stack section which is determined by that operator.

After applying the independent subquery method [8, 14] to (7) it takes the form (8).

\[
\begin{align*}
(wds(\text{"higher"}) & \text{ as } \$aux3, \\
wds(\text{"secondary"}) & \text{ as } \$aux2, \\
wds(\text{"vocational"}) & \text{ as } \$aux1).
\end{align*}
\]

\[
(emp \text{ as } \$emp \text{ where} \\
\quad \text{if } \$emp.\text{education} = \text{"vocational"} \text{ then } \$emp.\text{overtime} > \$aux1 \\
\quad \text{else if } \$emp.\text{education} = \text{"secondary"} \text{ then } \$emp.\text{overtime} > \$aux2 \\
\quad \text{else } \$emp.\text{overtime} > \$aux3)
\]

In (8) the weakly dependent subquery is evaluated only three times. The query (8) terminates optimization actions - no further transformations are possible by using the above methods. In consequence the query (8) will be converted into its XQuery counterpart:

\[
\begin{align*}
\text{let } \$aux3:= \\
\quad \text{avg(doc("company.xml")//emp[education="higher"]/overtime)} \\
\text{let } \$aux2:= \\
\quad \text{avg(doc("company.xml")//emp[education="secondary"]/overtime)} \\
\text{let } \$aux1:= \\
\quad \text{avg(doc("company.xml")//emp[education="vocational"]/overtime)}
\end{align*}
\]

\[
\begin{align*}
\text{for } \$emp \text{ in } \text{doc("company.xml")//emp} \\
\quad \text{where } \text{if}(\$emp/\text{education}="\text{vocational"}) \text{ then } \$emp/\text{overtime} > \$aux1 \\
\quad \text{else if}(\$emp/\text{education}="\text{secondary"}) \text{ then } \$emp/\text{overtime} > \$aux2 \\
\quad \text{else } \$emp/\text{overtime} > \$aux3
\end{align*}
\]

\[
\text{return } \$emp
\]

4. More general case

The examples in Chapter 3 presented only specific case of our transformation. The subquery (5) was weakly dependent from the most external operator. In general, however, the dependency of a subquery can be considered towards internal operators. The query (10) presents such a case.

\[
\begin{align*}
\text{for } \$dept \text{ in } \text{doc("company.xml")//dept} \\
\text{where every } \$emp \text{ in } \$dept/\text{employs/emp satisfies} \text{ emp/salary} > \\
\quad \text{avg(doc("company.xml")//emp[gender=$emp/\text{gender}]/salary)}
\end{align*}
\]

\[
\text{return } \$dept
The query (10) retrieves departments which each employee has the salary greater than the average salary calculated for all employees having his/her gender. The SBQL equivalent of the query (10) has the form (11):

\[
\text{dept as } \$\text{dept where } (\text{dept.employs.emp as } \$\text{emp} \forall \$\text{emp.salary} > 1 \text{ 2 3 3 3 3 3 4 4} \\
\text{avg}((\text{emp where gender=}$\text{emp.gender}$).salary)) 1 \text{ 4 4 3 5 5 4 4 4}
\]

Consider the following subquery of (11):

\[
\text{avg}((\text{emp where gender=}$\text{emp.gender}$).salary)
\]

The subquery (12) is weakly dependent only from the quantifier operator because it involves the name $\text{emp}$ in the expression $\text{emp.gender}$ that is bound in the scope opened by this operator and $\text{emp.gender}$ is typed by the enumeration. According to our rewriting rule \[5,6\] the proper conditional statement is put directly after the operator on which (12) depends:

\[
\text{(dept as } \$\text{dept) where } ((\text{dept . employs . emp as } \$\text{emp) \forall}) \text{ (if } (\text{emp . gender="male") then}} 3 \text{ 3 4 4} \\
(\text{emp . salary} > \text{avg}((\text{emp where gender="male"}).salary)) 3 \text{ 4 4 1 4 4 4 4 4} \\
\text{else } (\text{emp.salary} > \text{avg}((\text{emp where gender="female"}).salary)))) 3 \text{ 4 4 1 4 4 4 4 4}
\]

The query (13) illustrates the general rule of transforming a weakly dependent subquery. The rule can be applied both to external and internal operators. As a result of our rewriting the query (13) contains two subqueries

\[
\text{avg}((\text{emp where gender="male"}).salary)
\]

\[
\text{avg}((\text{emp where gender="female"}).salary)
\]

which are independent both from the quantifier and from the first where operator. Denote (14) by \text{wds}("male") and (15) by \text{wds}("female"). After applying the independent subquery method \[8, 14\], the query (13) takes the following form:

\[
(\text{wds("female") as } \$\text{aux1, wds("male") as } \$\text{aux2). (dept as } \$\text{dept) where } ((\text{dept.employs.emp as } \$\text{emp) \forall}) \text{ (if } (\text{emp.gender="male") then } (\text{emp.salary}>\$\text{aux2}) \\
\text{else } (\text{emp.salary}>\$\text{aux1)))}
\]
After converting (16) to the proper XQuery expression it takes the form (17):

\[
\begin{align*}
\text{let } &\text{$aux2:=avg(doc("company.xml")//emp[gender="male"])/salary) } \\
\text{let } &\text{$aux1:=avg(doc("company.xml")//emp[gender="female"])/salary) } \\
\text{for } &\text{$dept in doc("company.xml")//dept} \\
\text{where every } &\text{$emp in $dept/employs/emp satisfies} \\
\text{if($emp/gender="male") then $emp/salary > $aux2} \\
\text{else $emp/salary > $aux1} \\
\text{return } &\text{$dept}
\end{align*}
\]

(17)

5. Rewriting rule

The rewriting rule for queries involving weakly dependent subqueries can be formulated as follows. Let \( q \) be an XQuery expression of the form (18):

\[
\begin{align*}
\text{for } &\text{$i in doc(\text{uri})//q} \\
\text{where } &\text{$q1$} \\
\text{return } &\text{$i$}
\end{align*}
\]

(18)

where \( q2 \) has the form \( q2 = \alpha^\circ\text{wds($i/q3$)}^\beta; \alpha \) and \( \beta \) are some parts of \( q2 \) (maybe empty), \( \circ \) is a concatenation of strings, \( \text{wds($i/q3$)} \) is a weakly dependent subquery whose part \( $i/q3$ \) depends on the parent query only and is of the enumerated type \( ET = \{e_1,e_2,...,e_n\} \). Then the expression (18) is transformed into the following SBQL query:

\[
\begin{align*}
q_1 &\text{ as } $i where } q_2' \\
\text{where } &\text{$q2'=\alpha^\circ\text{wds($i/q3$)}^\beta'$ is an SBQL equivalent of $q2$. The query (19) is rewritten to the form (20) according to the rule presented in [5, 6].}
\end{align*}
\]

(19)

\[
\begin{align*}
q_1 &\text{ as } $i where if($i.q3=e_1) then \alpha''\text{wds($e_1$)}''\beta'' \\
&\text{else if($i.q3=e_2) then } \alpha''\text{wds($e_2$)}''\beta'' \\
&\text{else if($i.q3=e_3) then } \alpha''\text{wds($e_3$)}''\beta'' \\
&\text{......} \\
&\text{else if($i.q3=e_{n-1}) then } \alpha''\text{wds($e_{n-1}$)}''\beta'' \\
&\text{else } \alpha''\text{wds($e_{n}$)}''\beta''
\end{align*}
\]

(20)

After applying the independent subquery method [14] to (20) it takes the form (21):

\[
\begin{align*}
\text{(wds($e_1$))'}\text{group as } $aux_1, \text{wds($e_2$)}'\text{group as } $aux_2, \ldots, \\
\text{wds($e_n$)}'\text{group as } $aux_n). \\
(q_1 where if($i.q3=e_1) then \alpha'''\text{$aux_1$}'''\beta'' \\
&\text{else if($i.q3=e_2) then } \alpha'''\text{$aux_2$}'''\beta'' \\
&\text{else if($i.q3=e_3) then } \alpha'''\text{$aux_3$}'''\beta'' \\
&\text{......} \\
&\text{else if($i.q3=e_{n-1}) then } \alpha'''\text{$aux_{n-1}$}'''\beta'' \\
&\text{else } \alpha'''\text{$aux_n$}'''\beta''
\end{align*}
\]

(21)
The query (21) is converted into the following XQuery expression:

\[
\begin{align*}
&\text{let } \$aux_1 := \text{wds}(e_1) \\
&\text{let } \$aux_2 := \text{wds}(e_2) \\
&\quad \ldots \\
&\text{let } \$aux_n := \text{wds}(e_n) \\
&\text{for } \$i \text{ in } \text{doc(uri)//q} \\
&\quad \text{where if } ($i/q_1=e_1) \text{ then } \alpha^\$aux_1\beta \\
&\quad \quad \text{else if } ($i/q_1=e_2) \text{ then } \alpha^\$aux_2\beta \\
&\quad \quad \quad \text{else if } ($i/q_1=e_3) \text{ then } \alpha^\$aux_3\beta \\
&\quad \quad \quad \quad \ldots \\
&\quad \quad \quad \quad \text{else if } ($i/q_1=e_{n-1}) \text{ then } \alpha^\$aux_{n-1}\beta \\
&\quad \quad \quad \quad \quad \text{else } \alpha^\$aux_n\beta \\
&\text{return } \$i
\end{align*}
\]

(22)

6. Conclusions

We have adopted the optimization techniques used for SBQL to XQuery expressions. The approach consists in transforming an XQuery expression into its SBQL equivalent. The result of this operation is next rewritten according to the predefined rule. Finally, the optimized SBQL query is converted into an XQuery expression. The proposed optimization method was aimed at limiting the number of processing of a weakly dependent subquery to the number of enumerators of the enumerated type that the subquery depends on.

In the future we are going to express semantics of XQuery in terms of SBA. It allows to avoid converting XQuery expressions into SBQL ones. The SBQL semantics is based on three data structures: an object store, an environment stack, and a query result stack. It respects several principles such as total internal identification, orthogonal persistence, and compositionality. These features much simplify developing query optimization methods. For this reason it is worth to apply concepts related to the Stack-Based Approach for defining XQuery semantics.

REFERENCES


VISUALIZATION AS SUPPORT FOR WEB HONEYPOT DATA ANALYSIS

KRZYSZTOF CABAJ

Institute of Computer Science, Warsaw University of Technology

The paper presents methodologies associated with visualization, which supports data analysis. Analyzed data has been gathered by HoneyPot systems deployed in the network of Institute of Computer Science. Due to the vast amounts of data, the manual analysis was almost impossible and very impractical, also considering time constraints. Introduced visualization techniques and supporting filtering features are implemented in HPMS (HoneyPot Management System). The paper describes in details two introduced methodologies which support data analysis using both charts and graphs. The first one is used for the discovery of basic activities observed by HoneyPot. The second one is used for advanced analysis of machines used during attacks concerning PhpMyAdmin software.

Keywords: HoneyPot systems, visualization, data-mining, monitoring

1. Introduction

In the „Big data” days, the need for analysis of data gathered in vast amounts, by various information systems, cannot be questioned. For this purpose, multiple methods can be used, starting from the simplest one, implemented in each modern database system, which allows computation of average, minimal, maximal or count of given group of records, ending with the most advanced one that utilizes data mining techniques which could automatically group given data or discovered patterns appearing in the analyzed data. Although, described methods decrease the number of data presented, rapid discovery of most important information could be
very difficult and time consuming. Most often the new approach is used, which utilizes the visualization of data. The good visualization, even of vast amounts of data, allows easy discovery of most important relations or rules.

In the paper results and conclusions concerning conducted experiments associated with the analysis of HoneyPot logs are presented. The HoneyPot system is a kind of a trap for attackers, and data from such a system are gathered by the author since a few years now. Nowadays, vast amounts of such data are acquired, even dozen thousands events a month. For the analysis purpose, a dedicated system called HPMS (Honey Pot Management System) [1] was developed at the Institute of Computer Science, Warsaw University of Technology. The first implemented methods utilized simple data base functions which, for example, allow discovery of most active machines. During the following research, advanced data mining techniques that discover patterns are used, namely frequent sets and jumping emerging patterns, were designed. However, even using these methods a discovery of trends and associations between events was problematic. For this purpose the HPMS system has been extended with modules which allow data visualization. Currently, modules for plotting significant parameters during time and associations between various events using graphs have been implemented. The usage of such methods allows rapid discovery of time events, when new activity emerges or connections between them are discovered. In effect, the analytic work can be efficient, and reaction to new trends can be implemented.

The paper is organized as follows. The next, second, section describes HoneyPot systems, especially those which are used for gathering data for further analysis. The third section is associated with analyses that are supported by plots generated using HPMS systems. The fourth section describes the analysis of attacks directed to the PhpMyAdmin vulnerability using graphs. The last, fifth section concludes the conducted experiments.

2. HoneyPot systems

The role of the HoneyPot can be performed by any resource which can be used for observing hostile or unexpected activity. The only common feature of this resource is that it is not used for production purposes. The HoneyPot is usually a specialized machine or software; however, this role can be taken by a fake record in the data base or the account in the important computer system. Any access to the resource, for example, an attempt to read or login, is a sign of an unexpected activity. Historically, specially configured computers were used as the HoneyPot system. The configurations enabled various monitoring mechanisms that gathered as many data as possible concerning the attacker activity during an attack. For this purpose, logs from operating systems, logs from network devices placed between HoneyPot and Internet or even traces of all traffic directed to it can be used. Such a
solution was ideal for catching and tracking a human attacker, but it also had many drawbacks. The first and the most important one is associated with an additional risk. If the attacker detects and disables all the monitoring mechanism, the HoneyPot can be used for other hostile activity. Additionally, the initial deployment or cleaning the HoneyPot after a successful attack is very labor intensive. This kind of systems are called high interaction HoneyPots. In [2], the details concerning one of the first well documented development of the HoneyPot system and the description of further monitoring and tracing a real attacker can be found.

In the era of automatic threats, like worms, e-mail viruses or auto-rooters, dedicated high interaction HoneyPot systems used for gathering copies of new samples of malicious code are inefficient and very risky. After each infection the HoneyPot system must be cleaned. This process is relatively slow, even with the support of virtualization. A better solution for gathering information related to malware is the usage of low interaction HoneyPots. The low interaction HoneyPot is a dedicated software which imitates vulnerable services. Depending on the purpose, it can be very simple, for example, only listing for incoming connections and returning standard banners of a simulated service on the one hand. On the other hand, there are very complicated systems dedicated to downloading new samples of malware. This kind of low interaction HoneyPots simulates high level protocols in which vulnerabilities appear, incoming shellcode used by worm during vulnerability exploitation is emulated and the next stages of the activity of malware is downloaded. The most important low interaction HoneyPots are HoneyD [3], Nepenthes [4] and its successor Dionaea [5].

During our research on automatic threats, conducted at the Institute of Computer Science, only low interaction HoneyPots were used. Due to limitations of available systems, associated with very poor simulation of Web applications, a custom solution named WebHP was introduced. More details concerning its functionality are presented in papers [1, 6].

3. Analysis of HoneyPot data using charts

The WebHP HoneyPot system described in the previous section produces vast amounts of data that are later analyzed. For example, during the one and a half year the WebHP system receives more than 300 thousands of connections. Manual identification of new attacks or trends detection is almost impossible without dedicated software. Moreover, the analysis using only text representation of each connection is a tedious work. The most straightforward way to visualize this data is the usage of a chart. The Fig. 1 shows a sample chart from HPMS systems, which presents the activity observed by WebHP sensors during the year 2013. The X-axis
represents time and the Y-axis represents number of connections observed during each day.

Figure 1. Activity observed by WebHP sensors during the year 2013. Chart generated by HPMS software

At the first glance, a few very interesting time instants and time ranges can be identified. This information can be beneficial for a person which performs the analysis and points out the time range which should be investigated at first. In the presented Fig. 1, the sharp spike is associated with execution of security scanner. Both time ranges with high activity, the first between the end of June and middle of July and the second from October to the end of December, are related to automatic SPAM activity which placed marketing posts at the insecure webpage of the guest book. More details concerning this attack are presented in the article [1].

For more detailed analysis, the charts can be plotted for any tag associated with the connection. Tag definition is configured by the user, and later each connection which fulfill tags’ constraints is automatically marked by the given tag. Tag definition has a form of a logical expression which additionally can use any field appearing in the HTTP header. Fig. 2 shows sample logical expressions which are used for connection tagging.

header=['POST' configuration' AND (value ILIKE '42.php%' OR value ILIKE '43.php%' OR value ILIKE '44.php%' OR value ILIKE '45.php%' OR value ILIKE '124.php%')]

Figure 2. Sample tag definition used for automatic connection marking in the HPMS system

Using charts and tags provided by the HPMS system methodology for data analysis was introduced. At the first step the person which performs analysis for
a given time period observes the initial chart. Interesting features of the chart, for example, sharp spikes or time ranges which shows an observable rise in HoneyPot activity, are detected and later investigated in details. In many cases a detailed investigation reveals knowledge associated with detected feature in the chart, for example IP address of the offending machine or a repeated pattern in a field. In most cases using this information new tag can be created. In the next step a new chart, without previously detected transactions, is generated. The work is performed until there are no transactions which can be identified and marked.

The described methodology can be applied to data presented in the Fig. 1. Identified attacks lead to defining two tags that can be used for data filtering. Fig. 3 shows the activity of the deployed HoneyPot system in 2013 without these two attacks. As it can be seen, the usage of these two tags reduced the number of transactions almost ten times. In the new chart, other spikes can be observed. As was introduced in the methodology, these spikes should be analyzed and in effect new tags can be produced and additional knowledge associated with observed attacks can be gathered. Despite the fact that activity level in Fig. 3 compared to the Fig. 1 is one or two order of magnitude lower, methodology of analysis is almost the same.

Additionally, tags can be created using information gained from other source, for example, articles, reports or personal knowledge. Moreover, data mining methods implemented in the HPMS system are very beneficial, too. For this purpose the miner module, which discovers frequent sets in the recorded data, is used. The miner module treats each transaction as a set, and each transactions’ attributes as an item set. Detected frequent sets represent repeated activity. In many cases
cases such pattern can be easily converted to the tag definition and used later for marking new connections. Fig. 4 shows a sample pattern detected automatically using data mining techniques. The presented pattern is associated with ShellShock vulnerability [7] in the Linux program bash. Despite that bash is used locally by a logged user in most cases, some applications that use CGI interface and bash can be attacked too.

![Figure 4](image1)

**Figure 4.** A frequent set, automatically detected by HPMS miner module, which is later used for tag definition

![Figure 5](image2)

**Figure 5.** Transactions associated with ShellShock attack observed in 2014

When the tag is added, a new chart showing connections marked with recently added tag can be generated. In effect, after the detection of one attack, all similar ones can be detected too. The Fig. 5 shows connections marked with ShellShock
tag, which was produced using knowledge gained from discovered frequent sets. In the figure, additional six attacks can be spotted besides the dominating one.

4. Analysis of PhpMyAdmin attack using graphs

As was described in the previous section sample charts can be very beneficial for data analysis. However, not all data can be easily visualized in this manner, sometimes more complicated methods must be used. As an example of advanced graphical methods which visualize attackers and victims are presented. Research of this particular attack starts when at the end of February 2014 a completely new activity appears in the deployed HoneyPot. Manual analysis shows that this attack uses vulnerability in the PhpMyAdmin [8], utility popularly used by system administrators for MySQL database management using web interface. Rapidly new offenders start using this exploit and more and more new malware samples are received. Fig. 6 shows an activity associated with this attack by the deployed WebHP sensors.

![Figure 6. Automatically detected by HPMS miner module frequent set, which is later used for tag definition](image)

Before the introduction of the visualization method, some more details concerning this attack must be presented. The first attacker’s request looks like an innocent attempt to gather the file `/phpmyadmin/scripts/setup.php` (see Fig. 7). The file name is a little strange, but the web server responds with appropriate data. What is important to the attacker, is that the returned file contains a unique session token which is later used during an injection attack.
GET /phpmyadmin/scripts/setup.php HTTP/1.1
Host: 194.29.XX.YY
User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; MSIE 5.5; Windows NT 5.1) Opera 7.01 [en]

HTTP/1.1 200 OK
Date: Thu, 19 Jun 2014 08:51:35 GMT
Transfer-Encoding: chunked
Content-Type: text/html

24a7
<!DOCTYPE html>
<html xmlns="http://www.w3.org/1999/xhtml"><head>

... <a href="http://app/phpMyAdmin/setup/index.php?version_check=1&token=a60b82c06fa123c868288b29584d345">Check for latest version</a>
... </head>

<body>

... Figure 7. Initial message and response during the attack to the PMA application gathered by WebHP system

The second message which is sent from the attacker to the victim contains an exploit. In this case the attacker utilizes code injection error in PMA application that allows remote change of configuration object. In effect the attacker can download and execute any malicious code on the victim machine. Message containing the exploit, which was sent to the Honey Pot is presented in the Fig. 8.

POST /phpmyadmin/scripts/setup.php HTTP/1.1
Host: 194.29.XX.YY
User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; MSIE 5.5; Windows NT 5.1) Opera 7.01 [en]
Content-Type: application/x-www-form-urlencoded
Content-Length: 238

action=lay_navigation&eoltype=unix&token=a60b82c06fa123c868288b29584d345&configuration=a%3A1%3A%7Bi%3A0%3BO%3A10%3A"PMA_Config":1%3A%7Bs%3A6%3A"source"%3Bs%3A29%3A"ftp://37.59.AA.BB/pub/124.php"%7D}

... Figure 8. Message containing an exploit used during the attack to the PMA application gathered by WebHP system

The first marked part of the message is gathered in the initial phase of the attack – the session token. The second marked part (after decoding) contains the following text: “a:1:{i:0;O:10:"PMA_Config":1:{s:6:"source";s:29: "ftp://37.59.AA.BB/pub/124.php";}}”

21
This text contains the specially crafted PMA software configuration object which could be remotely changed due to an error in the application. In effect, the vulnerable software injects, in the currently executed instructions, the code downloaded from the provided malicious URL – file 124.php from IP 37.59.AA.BB. Because this exploit is sent to the low interaction HoneyPot, this attempt was only stored in the database and the malicious code did not execute. However, later manual analysis proved that the file 124.php contains malicious code which is used to remotely control the infected machine and to create a BotNet.

Sample analysis performed directly on data stored in the database could only find the most frequently used attackers, the used files or ftp servers from which malware was downloaded. This preliminary analysis rises interesting questions. If and how all those attackers, files and ftp servers are related? Is there any group of machines that works together? Do all such communities work in the same manner?

To answer all of these questions some visualization methods were introduced. Visualization utilizes simple undirected graphs, where a vertex represents an attacker (red), a filename (green) used by malware copy or an ftp server which hosts the malware (blue). The edge is added between a file and both an IP address – attacker and an ftp server. Conducted experiments prove the initial prediction that all the observed activity can be divided into independent groups. Fig. 9 shows a sample graph representing one of the detected groups.

![Graph presenting sample group of cooperating machines during attack](image-url)
Graphs generated automatically by HPMS system rapidly presents groups of attacks that are joined by the used file name or the ftp server. However, many other interesting behaviors can be easily spotted in these pictures. Fig. 9 shows that most activity is associated with one ftp server which hosts many files. Additionally, each file in most cases is used by three attackers. This probably is effect of redundancy introduced by attackers. Fig. 10 shows three other groups of offending machines.

Figure 10. Graphs presenting sample groups of offending machines. Two top groups are file oriented. The third at the bottom is attacker oriented.

The two graphs at the top of the Fig. 10 are file oriented, in contrast to the previous one presented in the Fig. 9 which is ftp server oriented. All attackers and ftp servers use the same file containing malware. What is interesting, in the graph at the left, all machines after infection download malware directly from the attacker IP. This is a sign that probably this is an automatic activity – some kind of the worm. The graph presented at the bottom of the Fig. 10 shows the attacker oriented graph. The central point of the graph is associated with an IP address of the offending machine. Used file has the same name, although lack of file name unification introduces two names – the only change is associated with the number of ‘/’ in the file name, which in actual connection has no effect. Like the graph at the right top of the figure 10, all the used ftp servers have login and password set. Used passwords are not so trivial which suggests that all servers are hacked, and store malware without owners knowledge and permission.
5. Conclusions

Deployment of the HoneyPot systems is just the beginning. Valuable knowledge can be gained by careful analysis of data which is gathered by the deployed HoneyPots. The real work with these systems begins at this point. However, currently most efforts concerning HoneyPot software development is associated with an introduction of new types of HoneyPots [9, 10] or hiding the existing ones [11]. The development of analytical software for HoneyPots is unfortunately neglected. The HPMS systems presented in this paper have been developed due to that no suitable analytical software exists. Huge amount of recorded data needs additional analytical functionalities. The visualization and filtering functionalities, initially implemented as support tools, rapidly confirmed their usefulness. As was presented in the previous paragraphs these simple methods can be very beneficial for people who perform analysis of data recorded by HoneyPot systems. The methodology described in the section 3, despite its simplicity, can be used with profound effects to analyze and characterize the activity observed by HoneyPot systems in the given time range.

Acknowledgments

Works associated with analysis of data obtained from HoneyPot system was funded by the Deans grant no. II/2014/GD/1 titled “Utilization of data mining techniques for network security data analysis”.

REFERENCES


THE DIAGNOSIS OF INFORMATION POTENTIAL
OF SELECTED BUSINESS PROCESS MODELLING NOTATIONS

RENATA GABRYELCZYK a), ARKADIUSZ JURCZUK b)

a) Faculty of Economic Science, University of Warsaw
b) Faculty of Management, Białystok University of Technology

The paper presents the findings of the ease of understanding tests for selected business process modelling notations as an element of diagnosing the notations information potential. The easiest-to-understand notation is identified, as well as the attributes determining the choice of this notation as the easiest to understand. Three notations used in business process modelling have been subjected to diagnosis: EPC, BPMS and BPMN. Based on the results of these analyses, recommendations have been formulated for organisations where process modelling requires the involvement of all employees and where process awareness needs to be developed. Using an intuitive notation may contribute to improved communication between users representing different professional profiles and translate into a higher effectiveness of organisational changes.

Keywords: business process modelling, modelling notations, modelling technique, BPMS, BPMN, EPC

1. Introduction

Over the past two decades, the concepts, methods and tools of the process approach have been continually evolving, leading to the formation of business process management (BPM) as a new field of organisational management in the contemporary economy. To implement BPM, an organisation needs to systemise the knowledge of its processes. This task involves the identification, documentation
and modelling of these processes. Process modelling can be applied to many areas. However, the most common is to formally define and document processes in order to fully understand them, enabling their continuous improvement and management. Business process modelling and management is constantly evolving, which is why, despite widespread literature on this subject, there are still ambiguities in terminology and a shortage of studies drawing attention to the differences between the terms, concepts, occurring standards and prerequisite choices of modelling notation.

This article aims to fill the gap in the cognitive information potential of the selected business process modelling notations. We can achieve this by evaluating the most commonly used graphical objects used to represent the actual business processes from the perspective of the user. The following modelling techniques were selected for analysis: EPC (Event-Driven Process-Chain), BPNM (Business Process Model and Notation) and the business notation of BPMS (Business Process Management System). Results of our desk research in this field is given in the first part of the paper.

Starting from the literature review we will base our analysis on the developed formal part of the evaluation process, conducting research on the information potential. In our research the ease of understanding of each of the selected business process modelling notation will be examined. The following research objectives were set for the study:

- to evaluate the ease of understanding the selected business process modelling notations as an element of diagnosing the information potential of the notation,
- to identify the notation perceived to be the easiest to understand,
- to identify the features which determine the notation to be selected as the easiest to understand.

The results of the study, presented in the second part of the paper, may provide crucial support in the choice of the methods and notations of business process modelling, depending on the area in which the modelling is applied. Recommendations for selecting the business process modelling notation at the operational level will be the practical purpose of the study.

2. Business Process Modelling

Business Process Modelling can be defined as a process of documenting business processes through a combination of text and graphic notation. In the context of business process management, it is most typically defined as a process used for mapping “the real world” (the as-is modelling), while being an active creation at the same time, which reflects the potential future states of the organisation or its processes, and suggests the potential directions of changes (the
to-be modelling) [11]. Process models help define processes and process interfaces, document processes, and present logical and chronological relations between process tasks, thereby enabling analyses, the assignment of agents, identification of information being transformed in the course of the process and information received as the process output.

Process models can be used as a basis for developing a performance measurement system applicable to processes, as well as to the whole organisation. In projects intended to enhance an organisation’s performance through improving its processes, Business Process Modelling requires the involvement of employees, external consultants and managers as well as the organisation’s rank and file from various departments. It is therefore essential for everyone to identical and comprehensible process modelling methods and tools. Furthermore, this requirement should prompt one to select an appropriate modelling notation which will:

- allow the necessary elements defining the process to be described at a given level of detail [1, 16, 17]
- contain graphic symbols and associated semantic rules comprehensible to all concerned, including the employees not professionally involved in process modelling [1, 14]
- facilitate communication between employees with different professional profiles
- enable the processes to be systemised logically and graphically within a framework concept, a not insignificant factor, as the models will be read and analysed by individuals who did not participate in the modelling directly [14, 15].

Process models that are comprehensible to employees will enable them to understand the processes performed throughout the organisation and will help them to view their tasks in the context of the entire value adding process. If a comprehensible notation is used in modelling, the employees may more readily engage in the identification, improvement and updating of the processes.

Business Process Modelling is a key element in the organisational change management and has many and varied applications, not solely limited to projects intended to develop a process-oriented organisation. Other important areas of the business process model application include arrangements preceding the selection or development of an IT system supporting business management (adjusting the system to the organisation, not vice-versa, a common language for IT and Business), designing workflow systems, documenting processes in the implementation of quality management systems, including ISO 9001 certification, and process benchmarking or Activity Based Management [15]. Therefore, selecting a notation gains significance in the context of the modelling objective and planned application of the business process model.
3. Assessment of modelling notation understandability: related works

Choosing an adequate Business Process Modelling Technique is one of the key issues to be considered at the stage of designing a business process model. Notation adequacy dictates that a model designed according to notation guidelines must have the adequate potential in terms of information and utility which meet the expectations of all its users. In the context of the Cognitive Fit theory [18], this potential can be viewed as the degree in which the notation fits the needs of the model user-supplier, the intended objective of model development and the form process information presentation. Hence, the information potential of the model depends on the properties deriving, primarily from its notation qualities, i.e. its structure, legibility, and ease of understanding [6]. Accordingly the assessment of the business process model potential should reflect the extent to which four key criteria are satisfied: ease of generation, ease of understanding, completeness, accuracy [10].

**Ease of generation** is the degree of ease associated with the design of business process models using a specific Business Process Modelling Technique (BPMT). In other words analysed BPMT is easy to use and it is easy to conceptualize a process using this approach [10]. **Ease of understanding** BPMT is the ability to design a process model easily understood by its users. A graphical representation of processes using BPMT are clear. **Technique completeness** means that process representation using this approach is complete and sufficiently detailed. **Accuracy** is the capacity in which BPMT allows the correct design of the business process models, and in accordance with business reality. High accuracy means that BPMT leads to accurate process representation [10]. A similar approach is presented by Recker&Mendling, who indicate that business process modelling technique/notation should allow for designing models that can be used as a basis for communication between users with different profiles (e.g. business vs. workflow analysts). Furthermore, the notation should be easy to comprehend, intuitive and should ensure interpretational flexibility of the model [13].

The findings of cross-sectional studies on methodologies used to assess the business process model comprehension reveal the consistency of approach that has been used. Table 1. contains examples of the objective and subjective measures used for assessing effectiveness of notations, taking into account the complexity of the purpose of studies presented in the article.

As independent variables, model notation and complexity were used in the research models discussed, while model comprehension/understanding and perceived ease of use understanding were taken as dependent variables. The dependence of variables was determined based on two groups of indicators: effectiveness and efficiency of the modelling notation used. Effectiveness of the models was generally measured by means of a comprehension test (e.g. multi-
choice comprehension questions). The efficiency assessment was based on relative and absolute measures representing the input required to understand the model (e.g. time needed to understand, number of correct answers/time of answers).

Table 1. Measures of Business Process Modelling Technique effectiveness

<table>
<thead>
<tr>
<th>References</th>
<th>Objective measures</th>
<th>Subjective measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct answers</td>
<td>Problem solving</td>
</tr>
<tr>
<td></td>
<td>on model content</td>
<td>based on model content</td>
</tr>
<tr>
<td>Bavota et al., 2011</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Figl &amp; Laue, 2011</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Fuller et al., 2010</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Kock et al., 2009</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>De Lucia et al., 2008</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Genero et al, 2008</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Mendling et al., 2007</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Recker et al., 2007</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Serrano et al., 2007</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Hardgrave et al., 1995</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Source: an analyses based on [5, 7, 8, 10, 12, 13]

4. The methodology of analysing the ease of understanding as an element of diagnosing the modelling notation information potential

The ease of understanding analysis was designed and carried out using a literature review and the business process modelling notation requirements for the area of application as a basis. The analysis constitutes an element of diagnosis of the modelling notation information potential.

According to the assumptions and the methodology accepted for the study, the analysis covered business process modelling notations based on an activity diagram and presenting a formal description of the process, with events and agents performing the activities addressed: EPC - Event-Driven Process Chain, BPMS - Business Process Management System and BPMN - Business Process Model and Notation. This is an unprecedented selection – the three notations not having been analysed before for their informational potential in terms of any of these criteria in such a combination as this. The notations selected have been used for modelling a process titled “Processing a freight forwarding order”. The structure of this process is compatible with the a generic model of a processing a freight order.
It initiates an external event "Forwarding order received". The result of the process is the preparation of shipping documents to implement a carriage service. An example of the analysed process model in chosen BPMT is shown in Figures 1–3. This modelling was performed by means of IT process modelling tools selected for the study, while the survey among respondents was conducted without indicating nor using these tools.

The method of Event-driven Process Chains (EPC) is a Modelling language used to describe business processes and workflows. EPC is the result of a collaborative project conducted by SAP AG and IDS Scheer AG in the years 1990-1992 [9]. This method was developed within the framework of Architecture of Integrated Information Systems (ARIS) in order to model business processes. EPC notation consists of events, functions and three types of connectors (logical AND, logical exclusive OR XOR and logical OR). According to EPC Modelling the assumption model consists of sequences of events triggering functions included in the business process. The whole process is triggered by the initial events [4].

In the case of the “Processing a freight forwarding order” process modelled in the study, EPC has been supplemented with elements of organisational structures (Fig. 1.) - in order to present the same elements of the process by means of each of the notations selected for analysis.

**Figure 1.** A fragment of the “Processing a freight forwarding order” process in EPC notation. Preparation in ARIS 9.7 Architect System
Business Process Management System (BPMS) is a framework for process management supporting a continuous process and total organisational improvement. BPMS is developed by BOC Information Technologies Consulting GmbH established in 1995 by Prof. Dimitris Karagiannis. The main idea of BPMS is to represent the dependencies between the core elements of an organisation: business process, product, organisational units, information technology and to make them controllable. According to BPMS approach modelling is done using the so-called model types which can be understood as "templates" for modelling processes, organisation units, roles, documents, etc [2]. The “Processing a freight forwarding order” process in BPMS notation is presented with elements corresponding to those in the remaining notations. A fragment of the process is shown in Fig. 2.

![Figure 2](image.png)

**Figure 2.** A fragment of the “Processing a freight forwarding order” process in BPMS notation. Preparation in ADONIS 5.1 Business Process Management Toolkit

Business Process Model and Notation (BPMN) is the global standard for process modelling, and was developed by Business Process Management Initiative (BPMI). Currently is maintained by the Object Management Group (OMG). According to their definition, BPMN is a graphical representation for specifying business processes in a business process model [3]. This notation allows both business modelling (basic level) and technical execution of processes (advanced level). At the basic level of complexity a model visually represents a business process flow (descriptive modelling). At the second level, the model gives possibility either to analyse the process performance using simulation tools or to create requirements for IT solutions (analytical modelling). The model complexity at the third level may deliver an executable code implemented as an application (executable modelling).

OMG assumption was to provide the same modelling notation understood by business analysts and technical developers [3]. Fig. 3. presents a fragment of the “Processing a freight forwarding order” process modelled using BPMN.
The survey was conducted among a group of the II-nd cycle university students (faculties: economics, business management), who were unaware of the process modelling techniques. Considering the respondents’ lack of experience in the modelling notations use, we assumed that their perception was similar to that of employees in organisations initiating business process modelling projects. The survey yielded 172 correctly completed questionnaires. The interviews were guided by the interviewers so as to allow all respondents the same response time.

The survey included two main tests of notation comprehension: a test diagnosing the subjective choice of the notation perceived as the easiest to understand; and, a test designed to identify the attributes determining the choice of the easiest to understand notation.

The general test of notation comprehension enabled the subjective assessment of understanding the business process flow and contents based on the model and a verification test. The detailed test of notation comprehension enabled subjective assessment of understanding a process fragment, with a particular focus on understanding logic gates. Furthermore, the detailed test included a verification component, which enabled objective assessment of respondents’ understanding of the notation.

For the purpose of comparative analysis, an ease-of-understanding indicator has been constructed for both tests of business process modelling notation understanding:

$$ Ease-of-understanding \ indicator \ for \ notation \ x = \frac{\sum x_c}{\sum x_s}, $$

where:

* $\sum x_c$ – the number of all respondents who answered the question verifying subjective perception of notation $x$ correctly,
* $\sum x_s$ – the number of all respondents who perceive notation $x$ as comprehensible.

**Figure 3.** A fragment of the “Processing a freight forwarding order” process in BPMN notation. Preparation in ADONIS CE 3.0: Free BPM Tool
The indicator gives a synthetic measure of notation understanding and allows the comparison of analysis results for all notations included in the study. The indicator can be used to analyse the results obtained in groups of respondents who have chosen the given notation as the easiest one to understand, as well as in the whole population of respondents, regardless of their choice of notation.

5. Research findings

According to the respondents’ subjective judgement, BPMS is easiest to understand – 61.6% of them chose this notation. Furthermore, the results of tests conducted in the respondent groups who have indicated the given notation as the easiest to understand, BPMS turned out to be the most comprehensible. The detailed comprehension turned out to be more difficult for respondents than the general comprehension test and more corroborative notation comprehension.

Table 2. Survey outcomes in respondent groups who have chosen the given notation as the easiest to understand

<table>
<thead>
<tr>
<th>Ease-of-understanding indicator for notation x</th>
<th>EPC</th>
<th>BPMS</th>
<th>BPMN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjective perceived ease of understanding a model</td>
<td>11.6%</td>
<td>61.6%</td>
<td>26.7%</td>
</tr>
<tr>
<td><strong>Test 1 – general comprehension test</strong> of ease of understanding of business process modelling notation</td>
<td>94.4%</td>
<td>96.1%</td>
<td>76.2%</td>
</tr>
<tr>
<td><strong>Test 2 – detailed comprehension test</strong> of ease of understanding of business process modelling notation</td>
<td>44.4%</td>
<td>57.0%</td>
<td>53.9%</td>
</tr>
</tbody>
</table>

An analysis similar to that summarised in Table 2. was conducted for the entire group of respondents, regardless of their choice of the notation easiest to understand. The outcomes support the findings of the survey in respondent groups who have chosen the given notation as the easiest to understand.

Percentage differences between the *ease-of-understanding indicator* value for each notation are insignificant. It is worth noting however that the biggest difference between Test 1 and Test 2 occurs for EPC, which may suggest that logical operators of this notation are the most difficult to understand. Even those respondents who indicated this notation as the easiest to understand made the most mistakes here showing that they did not grasp the process. The percentage difference between Test 1 and Test 2 is most negligible in the case of BPMN, which may prove that logical operators are presented in a more comprehensible
form here than in other notations, but only if the general process flow recorded in this notation is comprehensible.

Table 3. Survey outcomes for all respondents, regardless of their choice of the notation easiest to understand

<table>
<thead>
<tr>
<th></th>
<th>EPC</th>
<th>BPMS</th>
<th>BPMN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test 1 – general comprehension test</strong> of ease of understanding of business process modelling notation</td>
<td>89.2%</td>
<td>95.8%</td>
<td>75.3%</td>
</tr>
<tr>
<td><strong>Test 2 – detailed comprehension test</strong> of ease of understanding of business process modelling notation</td>
<td>38.2%</td>
<td>55.7%</td>
<td>53.9%</td>
</tr>
</tbody>
</table>

Table 4. summarises the analysis of notation perception and notation comprehension verification for all respondents. The percentage of responses given by respondents who declared their understanding of the given modelling notation in the total count of the survey participants (172 individuals) is the subjective measure. The percentage of correctly answered questions in the total count of respondents (172 individuals) is the objective measure. The reasoning is similar to previous analyses.

Table 4. Notation perception and verification of notation comprehension for all respondents

<table>
<thead>
<tr>
<th></th>
<th>EPC</th>
<th>BPMS</th>
<th>BPMN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test 1 – general comprehension test</strong></td>
<td>Subjective perceived understanding a model content /Count of all the survey participants</td>
<td>59.3%</td>
<td>96.5%</td>
</tr>
<tr>
<td></td>
<td>Correctly answered questions/Count of all the survey participants</td>
<td>52.9%</td>
<td>92.4%</td>
</tr>
<tr>
<td><strong>Test 2 – detailed comprehension test</strong></td>
<td>Subjective perceived understanding a model content /Count of all the survey participants</td>
<td>71.5%</td>
<td>91.9%</td>
</tr>
<tr>
<td></td>
<td>Correctly answered questions/Count of all the survey participants</td>
<td>27.3%</td>
<td>51.2%</td>
</tr>
</tbody>
</table>

The attributes determining respondents’ choice of notation were identified using the Kruskal-Wallis non-parametric method. The test was used to compare distributions of six qualitative variables (notation defined attributes) for three groups (notation types: EPC, BPMS, BPMN). The following dependent variables were used: number of graphic symbols, shape of graphic symbols, colour of
graphic symbols, use of pictographs, graphic way of describing business roles, and
graphic way of describing decision points.

Table 5. Kruskal-Wallis test outcomes for analysis of attributes determining the choice
of the easiest-to-understand notation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value of Kruskal-Wallis test</th>
<th>p-value (p &lt; 0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of the graphic symbols</td>
<td>2.4987</td>
<td>0.2867</td>
</tr>
<tr>
<td>Shape of the graphic symbols</td>
<td>7.9965</td>
<td>0.0183</td>
</tr>
<tr>
<td>Color of the graphic symbols</td>
<td>1.3883</td>
<td>0.4995</td>
</tr>
<tr>
<td>Use of pictographs</td>
<td>3.6253</td>
<td>0.1632</td>
</tr>
<tr>
<td>Graphic way of describing business roles</td>
<td>9.3722</td>
<td>0.0092</td>
</tr>
<tr>
<td>Graphic way of describing decision points</td>
<td>0.1596</td>
<td>0.9233</td>
</tr>
</tbody>
</table>

The dependent variables were measured by ordinal scale. A five-degree diagnosing scale was used. The results enabled assessment of the uniformity of respondents’ opinion distribution in respect of the factors determining their perception of notation comprehensibility. Table 5 summarises the outcomes of these computations. The computation outcomes prove that the shape of graphic symbols and the graphic way of describing business roles are the attributes determining the level of notation comprehension.

6. Conclusions and future research

The BPMS notation (Business Process Management System) was chosen subjectively as the easiest to understand by 61.6% of the respondents (26.7% BPMN, 11.6% EPC). The notation ease-of-understanding indicator is highest for BPMS in the respondent groups who chose this given notation as the easiest to understand, as well as regardless the easiest-to-understand notation choice. BPMS shows highest ease-of-understanding indicator values in the notation perception and notation comprehension verification analyses for of all respondents.

An analysis of the attributes determining the choice of the easiest-to-understand notation shows that the graphic way of describing business roles and the shapes of graphic symbols are most important. This information may be useful for those developing IT tools supporting business process management.

Based on the findings of the study, BPMS can be recommended as the most comprehensible notation for use in organisations where process modelling requires the involvement of all employees with their process awareness needing to be built as well as the “process-oriented” work style explained in the course of training sessions, or within the frames of process documentation projects for the purpose of
ISO certification. Understanding the organisation’s process models may enhance employees’ commitment, mitigate their resistance to changes and improve the effectiveness of company projects.

Our future research will focus on other aspects of diagnosing the information potential of notations and on comparative analyses of various groups of notation users.

REFERENCES

Arbeitspapier Nr. 72, Lehrstuhl für Wirtschaftsinformatik, Universität Hohenheim, 14-15.


ONLINE PERSONAL FINANCE MANAGEMENT APPLICATIONS

ViEra GafrikOvA\textsuperscript{a)}, WIESLAW SzCZESny\textsuperscript{b)}, ZDZISŁAW ODRzyGÓŻDŻ\textsuperscript{c)}

\textsuperscript{a)} Faculty of Management and Finance, University of Technology and Economics, Warsaw
\textsuperscript{b)} Faculty of Applied Informatics and Mathematics, Warsaw University of Life Sciences
\textsuperscript{c)} Faculty of Mathematics and Information Science, Warsaw University of Technology

The paper is devoted to internet applications supporting personal finance management (PFM). The main aim of the paper is to describe presumptions for development of PFM applications and to compare – from the user’s point of view – selected PFM websites in Polish, taking into consideration two categories: credibility and functionality. Within each category several factors were considered. The factors’ assessment has resulted in the creation of websites positioning.

Keywords: personal finance management, PFM websites, PFM internet applications, household budget

1. Introduction

Personal finance is a term that is increasingly appearing on radio, television, internet, etc. Also, various aspects of personal finance are increasingly becoming the subject of scientific research.

In the literature one can find different definitions of personal finance (called also household finance), for example, according to Business Dictionary [1] it is “Use of the principles and techniques of corporate finance in an individual's money affairs, especially the methods of allocation of financial resources. Its objective is financial security and independence so that an individual or a family can meet expected expenses and withstand monetary emergencies. It involves making prudent
financial decisions, budgeting, saving, investing, insurance, tax planning, retirement planning, and estate planning.

In the work [2, p.139] household finance is understood as economic phenomena occurring within a household that are directly associated with the collection and disbursement of money funds. In turn, the household’s financial management is interpreted as all household activities for the acquisition of funds and spending them in accordance with established household objectives. Some other definitions of personal finance are presented, for example, in the book [3, p.32].

Personal finance, therefore, includes a wide range of issues, inter alia, liquidity management, budgeting, capital raising, investing, saving, short-term and long-term planning, insurance and retirement security. Along with a dynamic development and internationalization of financial markets, increasing complexity and multiplicity of financial products, the development of information and communication technologies, the uncertainty in the labor market and pension reforms, personal finance management has become increasingly complex.

Examples of recent financial crises show that a lot of people cannot deal with their finances, which manifests itself, for example, as excessive indebtedness or investing in inappropriate, undiversified financial instruments.

The problem of insufficient financial literacy refers to people of many countries, including Poland, which is confirmed by results of studies. For example, the study "The Map of Financial and Insurance Education" [4, p.3], organized by the Polish Bank Association and the Polish Chamber of Insurance, says that many people do not understand simple economic phenomena and relationships between them, and hence, among others, they are not able to find themselves in an environment of free market economics. In the economic practice, a particular barrier is unsatisfactory financial education affecting the low level of people’s financial competence and the large area of exclusion in access to financial services.

Financial education as well as facilitation and assistance in the daily personal finance management is important for the proper functioning of people in the society. In the Western Europe and the United States of America this has been recognized for a long time. In Poland, young people get the basic financial and economic education in schools, though in a limited scope, as schools lack modern teaching tools and qualified teachers [4, p.4]. It is necessary to disseminate knowledge by other entities, among which the main role is played by central institutions such as the National Bank of Poland, the Polish Financial Supervision Authority, the Polish Bank Association, the Credit Information Bureau, foundations (the Kronenberg Foundation, Foundation for Youth Entrepreneurship), banks and credit unions. Specialized PFM tools for facilitating the daily personal finance management are offered by some banks and some non-financial companies.

The major purpose of this paper is to review and assess online PFM applications offered by non-banking companies and available for Polish-speaking users.
2. The cradle of PFM tools

Financial record keeping itself requires no special tools – a pencil and paper is enough. In the times of universal access to personal computers and the internet, the tedious tracking of income and expenses can be eased.

The most straightforward way to use computer would be working with spreadsheets. The advantage of such solutions is flexibility – a user himself determines a spreadsheet form, classifies data into categories, decides on the degree of detail, selects key performance indicators automatically updated, etc. The disadvantage of this approach is the necessity of possessing skills required for spreadsheets, the need to devote time to the template construction, to fill it with data and to adapt it for new purposes.

Keeping the household budget may be supported by specialized tools – so called PFM tools. These tools provide users with the information on the condition of their finances, the structure of expenditure and revenues, help create budgets, financial plans in the short and long term perspectives, categorize expenses, have extensive graphics and they also often allow the aggregation of financial information from various sources.

Personal finance desktop products emerged in the US in the 80s of the last century. They initially took the form of applications installed on the user’s computer and the information had to be manually entered into the program. Then the applications have been enhanced by the possibility to import a bank transactions history using internet banking files. Currently, web applications which are delivered primarily by non-financial institutions, but also by the banks themselves as a separate module in the online banking or mobile banking, become increasingly popular.

Internet analogues of PFM desktop products also have first appeared in the USA, where they become rather popular and where users can choose from a wide variety of these tools. US website Mint.com is considered an American benchmark. It also has inspired creators of similar applications in other countries. Mint.com was launched in September 2007 and quickly attracted more users than any other online solution. Over half a million people opened Mint.com accounts in its first year—making Mint.com the leading and fastest growing service of its kind. In 2009, Mint.com was acquired by Intuit for 170 million dollars. Currently, the service has more than 10 million users [5].

Mint.com cooperates with several thousands of financial institutions in the US and Canada. The user may see the real picture of his accounts: credit cards, investment, savings and credit accounts. Using the secure transfer of payment information, PFM systems have comprehensive data on user’s financial flows. Analytical modules of the system divide the stream of all transactions into categories, specifying the items of expenditure and summarizing them, specifying how much was spent on rent, transportation, food, clothing, restaurants, etc. Furthermore, they
show the outstanding balance of all loans and investments and interest both paid and received.

Advanced technology of Mint.com categorization was patented by the company. Mint.com slogans, not without reason, indicate saving money by users: the service reviews all expenditure and shows what money can be saved, doing it in a personalized manner. The system shows all expenses already made, recalls the need not to exceed budget, proposes efficient variants of the transaction – such messages come via e-mail. Also, reminders for payments due, information on a shrinkage of available funds and unusual expenses are sent by e-mail or SMS.

Particular attention is paid to safety. After all, users entrust to the system Mint.com not only information on their budget, but also on bank accounts passwords. Mint.com uses the same methods of cryptographic protection and physical security of information, as the leading US banks. The company has successfully passed the banking certification process.

Security is a key problem in the functioning of PFM website. Site Wesabe, one of the main competitors of Mint.com, which also required passwords for bank account access, ceased its activities in 2010, because of financial constraints – it was not able to provide its customers with an appropriate level of security.

Most PFM websites are free for users. Services make money on fees and commissions from banks, insurance companies, loan companies and other financial institutions, which are recommended to users. This is the American favorite business relations model, so called “win-win model” – all parties win.

A few years ago it seemed that PFM tools would become increasingly popular and would have rapidly growing number of users. Although in the United States, there is a number of PFM websites that have caught users’ attention and have achieved financial success, e.g. Mint.com, PageOnce (later Check and now Mint Bills) or Money Desktop, the consumer interest is not as high as expected. According to Aite Group’s 2012 survey [6] 58% of US consumers have not used and don’t plan to use a PFM platform. Only 27% of customers use PFM tools provided by banks or non-banking companies. Growth of PFM services is inhibited by a few factors, among them factors of cultural nature. The survey of Aite Group shows that only four in ten customers manage a household budget, categorize spending or forecast spending at least once a month. Only 13% of consumers want the help of financial institutions in their money management. So, what is the current state of PFM tools? The answer to this question might be the statement of Mr. Ryan Caldwell, chief executive officer of MoneyDesktop “We see PFM as helping someone fully understand and manage finance. This is a progressive game in its early stages” [7].
3. The overview of Polish language PFM applications

In Poland, personal finance management applications are offered by some banks and some companies, other than banks.

Several banks provide PFM applications (e.g. BPH, ING, Meritum, Millennium, T-Mobile Usługi Bankowe – former AliorSync). The main minus of the bank applications is that customers can manage only accounts maintained by the given bank, while the main plus is that banks are able to ensure the appropriate level of security. On the other hand, PFM applications offered by non-banking companies, can be supplied by different institutions’ accounts. Among offered services are: online applications (mostly free) and programs downloaded from the internet to be executed on a PC, free of charge or paid (e.g. Skarbnik, Menedżer Finansów, e-Portfel, ShuCash), some of them can be purchased also in a box version as well.

This section analyzes selected online PFM tools in Polish offered by non-banking companies. Due to the volume of the paper, only the short characteristics of websites, which include these tools, are provided below.

- **Budżet domowy [8]**
  Budżet domowy is the first PFM tool in Polish, created in 2005. The website owner is residing in London. Both the site and the application are protected by the copyright law of the UK and international agreements. The homepage is interesting, with the dynamic movable design, though it seems overloaded with the information. The website contains tips for saving money in different areas and the information on basic banking products. A demo version of the application is available.
  There are three ways of downloading data into application: manually, importing internet banking files (several banks, formats CSF, OFX, QIF), automatic access to online user’s banking accounts (two banks mBank, PKO BP). Expense and income categories are predefined with the possibility to create users’ own categories. Budżet domowy enables the budgeting in categories and provides the basic graphical presentation of the budget execution. The export of data in CSV format is possible, but this function is for pay, while all other functions of the application are free. According to the information on the site, all communication is encrypted in the same way that communication with the bank is, but more detailed information is not provided.

- **e-Grosik.eu [9]**
  Using the service is free. The home page is moderately attractive. It contains a link to the Open Finance, which enables to compare mortgage offers from over 20 banks. The main page contains also non-financial advertisements, e.g. about medical cosmetics, which lowers the professionalism of this website.
Once the account is created, a user can manually enter his expenses and incomes (this is the only possibility to provide data). The application proposes its own categories and subcategories of incomes and expenses, but a user may create his own categories and subcategories. It also creates several kinds of reports for selected period in the tabular form, and some of them also in the graphical form.

- **Figgy.pl** [10]
  The owner of the website is Devesoft Software Development. The home page is encouraging, although there is the information, that the site is in the beta version. A user defines his own income and expense categories. Transactions are downloaded manually and may be exported into CSV or PDF files on the daily basis. A user may define a list of standing orders and a list of user’s favourite stores. The application offers a simple graphics. All communication with the site is performed using a secure 128 bit encryption (Thawte certificate). The service is free.

- **Kontomierz** [11]
  The website was created in 2008. The site owner is the company Kontomierz, Ltd. The design of home page is attractive, inviting and the space is used efficiently. The demo version is available. Currently all application functions are free. The company plans to provide paid accounts with additional elements, in the future. However, the basic functionality will be free. Kontomierz possesses security certificate of Verisign Trusted Seal, the most reputable company issuing certificates for web applications.
  At present, Kontomierz imports current account transactions from over 30 banks, either by automatic import or by means of transactions files from internet banking accounts. The manual import of data is available. It was mainly intended for cash payments, which can be also entered to the application by sending SMS to the special phone number. Export of data in the CSV form is available.
  Predefined expense and income categories are built into the application. A user may define monthly budget limits for any expense category. The application follows these limits and informs a user about exceeding them. Results are presented also in the graphical form. The application sends automatically e-mail reminders, e.g. about lack of the activity on the Kontomierz account or about suspiciously large transactions. A mobile version of the application is also offered.

- **MojeGrosze.pl** [12]
  The website was created in 2008. The main page is well organized and easy to use. A demo version is available.
  The data can be imputed manually or by the import of transactions from the internet accounts. A user defines his own categories. A simple graphics is available. The website has not been developed for some time: the representative of the company confirmed unofficially that work on this website had stopped for the time being.
• MoneyZoom [13]
The main page of the website is elegant, spacious, well organized and therefore user friendly. The website contains a financial guide with advices concerning money management in different categories. The price comparison tool of the Comperia.pl, the strategic partner of MoneyZoom, is available from the site.

The manual introduction of data is the only possible way. The tool proposes its own income and expense categories and subcategories. A user can also create his own categories and subcategories. It is possible to introduce spending limits and plan future expenses. A graphical analysis is possible. MoneyZoom.pl has implemented the latest data encryption standard used for banking sites. It created in 2010 the first Polish PFM application on the iPhone. This allows users to control their budget, wherever they are.

• Payzzer [14]
Payzzer is a new service – it was created in 2014 and replaced the service Finser which was created in 2010. The home page is attractive and inviting, with a demo version. The website is accessible in Polish and English and includes the extensive FAQ section. Only the manual data entering is possible. A user creates his own categories. A graphical analysis is available. Payzzer has its mobile version and the application operates on Android platform as well.

4. Positioning of PFM websites

Websites were assessed from two points of view: credibility and functionality, considering several factors. Factors were evaluated by means of experts’ estimates on a few points scale. Then estimates were summed up to obtain a composite estimate of the given website – on this base, websites positioning was performed.

4.1. Factors taken into account in assessing credibility

• Security
Security issues must be treated very seriously as websites have access to users’ financial data. Data security requirements should be at a level comparable to the requirements of banks. Security was assessed on the basis of the information disclosed on the website pages as follows: 2 points – for the highest quality security certificate (displays green padlock and website owner’s company name in the green address bar, https:// at the beginning of the domain name - signs that a site can be trusted); 0 point – for lack of information on the manner of data encryption, 1 point – in other cases.
**Website owner**

Basic information on the owner should be placed on pages of the website – this is the easiest way to show that there is a legitimate and real company behind the website. Users trust the site that is owned by a firm with a good reputation. Coordinates of the company enable a user to check further information on company’s activities which he is interested in. As a quoted joint stock company should provide most detailed information on its business activities, evaluation was the following: 2 points - quoted joint stock company; 1 point – other company registered in Poland; 0,5 point – other company registered abroad; 0 point - lack of information on the owner.

**Contact data**

Users should be helped to solve their problems while using a service on a website as quickly as possible, therefore: 2 points - contact by phone and e-mail, 1 point - contact by e-mail only, 0 point - no contact.

**Design**

Websites that are more attractive create a greater feeling of trustworthiness and professionalism in customers. Two cases were taken into consideration: an attractive website (2 points) and not very attractive, modest website (1 point).

**Updating**

People assign more credibility to sites that show they have been recently updated or reviewed. The information on the site should be current and well maintained, hence: 2 points if the information on the site is updated regularly; 1 point - if the site information is basically updated; 0 point - if the website contains the information which has not been updated for several years.

**User’s responsibility**

Information on the extent of users’ responsibility should be precise, detailed and easy to understand, consequently: 2 points for clear and comprehensive information on user’s responsibility, 1 point for general information, 0 point for no or very limited information. The results of websites credibility assessment are presented in the Table 1.
Table 1. Websites positioning according to their credibility

<table>
<thead>
<tr>
<th>Website</th>
<th>Assessed factors</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Security</td>
<td>Website owner</td>
</tr>
<tr>
<td>Budżet domowy</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>eGrosik.eu</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Figgy.pl</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Kontomierz</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>MojeGrosze.pl</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MoneyZoom</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Payzer</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

4.2. Factors taken into account in assessing functionality

- **Data downloading**
  Basically, there are three ways to load data into an application. The first way is to load data manually (1 point). The second possibility is to load data by importing a file of transactions from internet banking accounts, e.g. a .xls, .csv, .tsv file (2 points). It might be a good solution for consumers that are not comfortable entering all of their bank information into PFM website. The third way is to import data by means of automatic access to online user’s banking records – to take advantage of this possibility a user needs to enter his banking ID and banking password. In the third situation, 1 point if an application can connect less than 10 banks or 2 points if the automatic import from more than 10 banks is possible. So, if an application offers all three possibilities of the data downloading, points are summed up.

- **Data export**
  In certain situations, the possibility to export data from PFM application to other program that has an import facility (e.g. Excel or Access) may be useful for a user: 0 point in case of a lack of such possibility, 1 point if import facility is available in the limited range, 2 points if an application has the data export facility.

- **Income and expense categorization**
  Tracking user’s expenses requires breaking down all expenses into appropriate categories. It helps user to learn about his spending tendencies and to control them. Depending on the manner of categorization of income and expenses, the scores are the following: 0 point if the application doesn’t propose a categorization, 1 point – a simple categorization, 2 points – a more advanced classification.

- **Graphical visualization**
  Graphical visualization of the analysis results is efficient, it helps users to look at large amount of data quickly. It helps also to achieve deeper insight into user’s
spending and to discover better understanding of spending. Points were assigned depending on the complexity level of the graphical display: 0 point for the lack of graphics, 1 point for the simple graphics, 2 points for the more advanced graphics.

- **Budget planner**

An application should have the possibility to budget expenses and income and then to monitor and control spending in order to keep it in line with planned budget limits. Points were determined in the following manner: 0 point for no possibility of budgeting, 1 point for the simple budgeting and 2 points for the advanced budgeting, with reminders sent automatically to a user, for example in the case of exceeding limits.

- **Financial guide and price comparison tool**

Some websites contain a financial guide where a user can find advices concerning money management in different categories such as “house”, “car”, “education”, “savings”, “food”, “entertainment”, etc. A price comparison tool, in turn, enables to compare prices of different financial products. The points are the following: 0 point for lack of such facilities, 1 point for a financial guide, 1 point for a price comparison tool. So, if there are both facilities available on the site, points are summed up.

### Table 2. Websites positioning according to their functionality

<table>
<thead>
<tr>
<th>Website</th>
<th>Data downloading</th>
<th>Data export</th>
<th>Income/expense categorization</th>
<th>Graphical visualization</th>
<th>Budget planner</th>
<th>Financial guide and price comparison</th>
<th>Mobile PFM solution</th>
<th>Composite estimate</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Budżet domowy</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>eGrosi\k.eu</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Figgy.pl</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Kontomierz</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>15</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>MrGrosze.pl</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>MoneyZoom</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>10</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Payzzer</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

- **Mobile PFM solution**

Some applications have versions on i-Phones or operate on Android platform. This allows users to track their financial data, wherever they are. If an application has a mobile version – 1 point, 0 point in the opposite case. Table 2 shows the websites positioning according to their functionality.
4.3. Final ranking

Summing up the estimates for the credibility and functionality, the final ranking of the investigated websites was obtained:

Table 3. Websites final ranking

<table>
<thead>
<tr>
<th>Website</th>
<th>Kontomierz</th>
<th>MoneyZoom</th>
<th>Budżet domowy</th>
<th>Payzzer</th>
<th>Figgy.pl</th>
<th>MojeGrosze.pl</th>
<th>MojeGrosik.eu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total estimate</td>
<td>26</td>
<td>21</td>
<td>16.5</td>
<td>14</td>
<td>9</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>Ranking</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

The research has shown that the current offer of online PFM applications in Polish is rather modest and until now customers are moderately interested in them – for example, MoneyZoom informs that it has 30,000 users.

The investigation has revealed that only two of the investigated websites (Kontomierz, MoneyZoom) provide the highest level of security, which is the basic condition for winning the confidence of the clients. These two sites remain – in our opinion - the key players on the Polish market. Some websites have stopped to develop or appear not to care about users (e.g. MojeGrosze.pl, Budżetdomowy.pl), as indicated – among others – by information on the site that has not been updated for a long time, posts on the discussion forum that are several years old or the lack of interest in solving users problems: our experience shows that receiving answers to user’s questions requires several e-mails and number of phone calls. Figgy.pl and e-Grosik propose the simple PFM tools and moreover Figgy.pl has functioned in so called beta-version for quite long time. Payzzer was created in the last year, so the time will verify its functioning.

5. Conclusions

Polish online PFM tools began to appear mainly in 2007-2009. Since that time the market has changed considerably: PFM websites which remain on the market develop at different rates, some have ceased their activities (for example 5groszy.pl, Benefi.pl and more recently, in December 2014, iFin24.pl), some have changed their original purpose (Moje.oszczędności.pl, domain currently listed for sale) and some new have arisen (e.g. Payzzer).

In the beginnings of online PFM tools in Poland, it seemed that some of them would repeat the success of American PFM web tools, naturally at a scale appropriate to Poland. Time has verified these expectations. There are many reasons for this – below we list a few of them.

The first group of reasons is connected with the economic performance of companies offering PFM tools. The first Polish PFM applications were mostly
created as a result of small firms work. Taking into consideration all costs, while applications were supposed to be free for customers, it was rather difficult to raise funds for further development and for ensuring the appropriate level of security.

The second group of reasons is related to the different financial behavior of consumers in the US and Poland, influenced by tradition and the different economic situation of both countries. Below are some differences.

In the United States using non-cash payments (checks, cards, money orders) has been common for many years, and then the expenses control is more difficult than in the case of using cash. The share of cash payments in the total volume of payments was 26.8% in 2012 [15, p.26], while cash remains the most popular payment method of consumers in Poland – 64% of the total volume of payments, and 82% of the total number of payments [16, p.257]. Further, US banks charge high penalty fees for overdraft or non-sufficient funds on accounts. The average fees of this kind, on accounts that had one or more overdraft transactions in 2011, amounted to $225 [17, p.5]. Related charges vary widely, depending on the bank. For example, the fee for overdraft at Bank of America is $35 in accordance with the current tariff of fees and charges. To avoid costly overdraft charges, Americans were used to keep track of their spending already before the era of PFM applications.

PFM services help not only to track expenses, but also help to save money. A habit of saving money, avoiding unnecessary spending and maximizing profit seem to be features typical for Americans. Also, results of studies confirm the propensity for savings: consumer ownership of saving accounts was 75.8% in 2012 [15, p.40]. In Poland, only 40.8% of households have savings, from which 66.6% in the form of saving deposits in PLN (data for 2013, [18, pp. 76-77]). In addition, the report on economic knowledge of Poles, published jointly by the Institute of Freedom and Raiffeisen Polbank [19, p.36], indicates that Polish people do not attach much importance to financial matters, do not strive to maximize their individual profits and are reluctant to make calculations.

The third group of reasons is related to the long US tradition regarding all personal finance issues. Personal finances are under investigation for almost 100 years (since the 30s of the last century, when the US financial market had been created), and the American experience with the personal finance management tools lasted now 20-30 years. Personal finance is separate, full-fledged scientific research direction. It is taught at American universities as a separate course, or as a part of general finance courses. Academic curriculum of personal finance is supported by adequate textbooks. Theoretical foundations of this field of finance and its development perspectives are discussed in the renowned scientific journals.

In Poland, the personal finance has been subject of the scientific research for about two decades and personal finance management courses appeared at universities only in the recent years.
Along with the increasing globalization of world markets and the further socio-economic development of Poland, above mentioned differences should be decreased gradually over time.

To sum up, PFM online tools are currently in their beginning stadium of development. The further progress of IT technology, the increasing consumers' financial literacy and the expansion of the range of users by the new generation for whom the internet is the natural environment, will lead to the improvement and perfection of PFM tools and to their more common usage.

REFERENCES


The advantages of Bayesian methods over classical methods in the context of credible intervals

Wioletta Grzenda

Institute of Statistics and Demography, Warsaw School of Economics (SGH)

The growing computational power of modern computer systems enables the efficient execution of algorithms. This is particularly important in Bayesian statistics, in which, nowadays, the key role is played by Markov Chain Monte Carlo methods. The primary objective of this work is to show the benefits arising from the use of Bayesian inference, especially confidence intervals in the context of logistic regression. The empirical analysis is based on "Household budgets" survey of Central Statistical Office. In this paper the unemployment among people over 55 will be investigated.

Keywords: Bayesian inference, confidence intervals, MCMC, logistic regression, unemployment

1. Introduction

Bayesian credible intervals, also known as credible regions (credible sets), play a similar role to classical credible intervals, but the philosophy of their composition and interpretation are quite different. The Bayesian approach gives the possibility of incorporating additional information that is external to the sample by prior distributions [8, 15]. This additional information may improve accuracy and credibility of estimations. The credible regions incorporate this prior information, while frequentist confidence intervals are based only on the sample data.
In frequentist methods, the parameters of a model are unknown, but fixed constants. Therefore, for a given sample, it cannot be specified whether the unknown value of this parameter is covered by this interval, or not. The randomness of this interval is based on the fact that for different samples we can obtain different realizations. The probability that the unknown parameter is within the confidence interval is either 0 or 1. However, in a long series of observed samples, the frequency of intervals, including true value of this parameter asymptotically equals $1 - \alpha$ (for example [7]).

Moreover, the interpretation of classical confidence interval may be at times senseless; for example when only one sample exists and additional samples cannot be gathered. Then it is Bayesian approach that gives a reasonable interpretation and classical confidence interval should not be applied.

In Bayesian statistics, the estimated parameter is a random variable. Then the credible regions may be exactly estimated with a probability level given a priori. The Bayesian credible region estimated from the current sample contains the estimated parameter with the given probability. This interpretation seems intuitive and frequentist confidence intervals are thus misinterpreted as Bayesian intervals [3].

For small sample surveys, Bayesian analysis is more accurate as asymptotic approximation is not used [1]. However, it is worth mentioning that the impact of a priori distribution on a posterior distribution may be more significant.

The empirical examples presented in this paper refer to the analysis of unemployment among older people. There is no commonly accepted age over which an individual enters this age group [4]; usually these are employees aged 50 and older or aged 55 and older. In this paper, people aged 55 and older have been investigated as the retirement age has been raised recently. The chances of finding a job depend mainly on the situation on the labour market and some demographic and socio-economic characteristics of an individual. In unemployment studies, logit models have been most frequently used (for example [5], [12]). In this paper, the Bayesian logistic model has been used to analyse the impact of different characteristics on one’s chances to find a job.

2. The definition of Bayesian confidence interval

Let $\Theta$ be the parameter space, $C$ be subspace of this space, $C \subset \Theta$. Moreover let

$$p(\theta | x) \propto p(\theta) p(x | \theta)$$

be posterior distribution of $\theta$. $\theta$ can be a parameter vector, as will be the case in next part of this paper.
Then the probability that parameter $\theta$ is in the space $C$ is given by

$$P(\theta \in C | x) = \int_C p(\theta | x) d\theta = 1 - \alpha, \ 0 < \alpha < 1. \quad (2)$$

This probability is the degree of belief that unknown parameter $\theta$ is in the space $C$, on condition that observations are available and we have a priori knowledge on this parameter. The region $C$ is called Bayesian credible region. Moreover, there are many possible credible regions for a given probability level $\alpha$, the smallest of them has been defined by formulas 3 and 4. It has been called the highest posterior density region (HPD).

If for the subset $C^*$ the parameter space $\Theta$, $C^* \subset \Theta$ holds

$$P(\theta \in C^* | x) = 1 - \alpha, \ 0 < \alpha < 1 \quad (3)$$

and for every $\theta_1 \in C^*$ and $\theta_2 \notin C^*$ holds

$$p(\theta_1 | x) \geq p(\theta_2 | x). \quad (4)$$

then $C^*$ is called $100(1 - \alpha)%$ the highest posterior density region.

The highest posterior density region is the region for which the minimum density of any point within that region is equal to or larger than the density of any point outside that region. This property refers only to Bayesian credible intervals and does not hold true in classical statistics [2, 10, 13].

3. Bayesian logistic regression

The logistic regression models explicate the relationship between a dependent binary or dichotomous variable and one or more independent variables [16]. In this paper, the binomial logistic regression will be investigated. Let $(X_1, \ldots, X_p)$ be a vector of independent variables, and $\beta = (\beta_1, \ldots, \beta_p)$ a vector of regression coefficients. Moreover, let $\pi$ be the probability of success i.e. of obtaining by a dependent variable one of two possible values $y_i = 1$ or $y_i = 0$ ($i = 1, \ldots, n$), on condition that independent variables have given values.

Let

$$\text{logit}(\pi) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p. \quad (5)$$

Then classical binomial regression model is given by

$$\pi = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p}}. \quad (6)$$
The likelihood function for a data set of \( n \) observations is

\[
L(\beta; y) = \prod_{i=1}^{n} \left( \frac{e^{\beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}}}{1 + e^{\beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}}} \right)^{y_i} \left( 1 + e^{\beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}} \right)^{(1-y_i)}
\]  

(7)

In this paper the Bayesian approach to logistic regression has been investigated [1, 6, 8]. Bayesian method needs an appropriately defined prior distribution. In the case of regression models, for regression coefficients \( \beta \) we choose \( p \)-dimensional normal prior distribution \( N_p(\mu_0, \Sigma_0) \), where \( \mu_0 \) denotes the prior mean vector, and \( \Sigma_0 \) denotes the prior covariance matrix. For \( j = 1, \ldots, p \)

\[
\beta_j \sim N(\mu_j, \sigma_j^2).
\]  

(8)

Then the posterior distribution by using Bayes' theorem is given by

\[
p(\beta | y) = \prod_{i=1}^{n} \left( \frac{e^{\beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}}}{1 + e^{\beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}}} \right)^{y_i} \left( 1 + e^{\beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}} \right)^{(1-y_i)}
\times \prod_{j=1}^{p} \frac{1}{\sqrt{2\pi\sigma_j}} \exp \left( -\frac{1}{2\sigma_j^2} (\beta_j - \mu_j)^2 \right)
\]  

(9)

In Bayesian approach, inference from any element of parameter vector \( \beta = (\beta_1, \ldots, \beta_p) \) is held on the posterior marginal distribution. Then the distribution is obtained by integrating the remaining part of posterior distribution. The Bayesian confidence intervals for the elements of parameter vector \( \beta \) are calculated by formulae presented in the second part of this paper.

In practice, for models with a large number of parameters, the simulation methods are used to generate samples from an arbitrary posterior distribution. Currently Markov Chain Monte Carlo Methods (MCMC) are used in estimations [17]. The Markov Chain Monte Carlo Methods are based on ergodic Markov chain, which over time converges into the stationary distribution. In Bayesian statistics, this stationary distribution is called posterior distribution. In this paper ARMS algorithm (Adaptive Rejection Metropolis Sampling Algorithm) has been used [11].

4. Empirical example

In this paper, a data set from the survey of Central Statistical Office of Poland – “Household budgets in 2009” and “Household budgets in 2010” has been used. In the first survey, 37302 households including 108038 respondents have been
examined and in the second one 37412 households with 107967 respondents. Professionally active individuals are people aged 15 and older, either employed or unemployed. In line with the aim of this research, only professionally active people aged 55 and older have been taken into consideration. Therefore, 4420 respondents have been chosen for analysis with the unemployment rate 5.75\% for 2009 and 8.39\% for 2010. By unemployed we mean people, who were looking for a job and were ready to take a job ‘this or the next week’. Farmers, gardeners, foresters and fishermen have not been investigated in this survey.

Table 1. The independent variables of models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Levels</th>
<th>Percent</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>man</td>
<td>1</td>
<td>65.63</td>
<td>64.22</td>
</tr>
<tr>
<td></td>
<td>woman</td>
<td>2</td>
<td>34.37</td>
<td>35.78</td>
</tr>
<tr>
<td>Marital status</td>
<td>unmarried, separated or divorced,</td>
<td>1</td>
<td>21.04</td>
<td>20.37</td>
</tr>
<tr>
<td></td>
<td>a widower, a widow</td>
<td>2</td>
<td>78.96</td>
<td>79.63</td>
</tr>
<tr>
<td>Education status</td>
<td>higher</td>
<td>1</td>
<td>21.61</td>
<td>17.49</td>
</tr>
<tr>
<td></td>
<td>post-secondary</td>
<td>2</td>
<td>2.87</td>
<td>6.47</td>
</tr>
<tr>
<td></td>
<td>secondary professional</td>
<td>3</td>
<td>27.01</td>
<td>25.21</td>
</tr>
<tr>
<td></td>
<td>secondary general</td>
<td>4</td>
<td>6.74</td>
<td>7.37</td>
</tr>
<tr>
<td></td>
<td>basic vocational</td>
<td>5</td>
<td>29.68</td>
<td>31.37</td>
</tr>
<tr>
<td></td>
<td>primary school</td>
<td>6</td>
<td>12.08</td>
<td>12.08</td>
</tr>
<tr>
<td>Region of Poland</td>
<td>central (łódzkie, mazowieckie)</td>
<td>1</td>
<td>23.87</td>
<td>23.45</td>
</tr>
<tr>
<td></td>
<td>south (małopolskie, śląskie)</td>
<td>2</td>
<td>18.17</td>
<td>20.23</td>
</tr>
<tr>
<td></td>
<td>east (lubelskie, podkarpackie, święto-</td>
<td>3</td>
<td>13.05</td>
<td>13.37</td>
</tr>
<tr>
<td></td>
<td>krzyskie, podlaskie)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>northwest (wielkopolskie, zachodnio-</td>
<td>4</td>
<td>17.01</td>
<td>16.43</td>
</tr>
<tr>
<td></td>
<td>pomorskie, lubuskie)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>southwest (dolnośląskie, opolskie)</td>
<td>5</td>
<td>12.35</td>
<td>11.82</td>
</tr>
<tr>
<td></td>
<td>north (kujawsko-pomorskie, warmińsko-</td>
<td>6</td>
<td>15.54</td>
<td>14.70</td>
</tr>
<tr>
<td></td>
<td>mazurskie, pomorskie)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Place of living</td>
<td>city of 100 thousand residents and more</td>
<td>1</td>
<td>37.49</td>
<td>35.21</td>
</tr>
<tr>
<td></td>
<td>city below of 100 thousand residents</td>
<td>2</td>
<td>30.00</td>
<td>31.46</td>
</tr>
<tr>
<td></td>
<td>country</td>
<td>3</td>
<td>32.51</td>
<td>33.33</td>
</tr>
<tr>
<td>Age group</td>
<td>60 years old and older</td>
<td>1</td>
<td>27.40</td>
<td>27.37</td>
</tr>
<tr>
<td></td>
<td>less than 60 year</td>
<td>2</td>
<td>72.60</td>
<td>72.63</td>
</tr>
</tbody>
</table>

*Source: own analysis of the data "Household budgets" 2009 and 2010*

The most frequently reported factors related to unemployment are: sex, marital status, education status, age and kind of previous job [5]. For this survey, based on initial modelling, the following independent variables have been chosen: sex,
marital status, education status, region of Poland (EUROSTAT) where a respondent lives, place of living and age. In table 1 the analysed characteristics of professionally active people have been presented.

Estimation and verification of all the models has been performed using SAS system. In the model, the number of burn-in iterations has been set to 2000 and the number of iterations after burn-in has been set to 10000.

In this paper, three models with non-informative and informative prior distributions have been estimated [14, 15]. The first model for data from 2009 has been estimated with non-informative prior distributions. The second model for data from 2010 has been estimated with non-informative prior distributions, too. The third model for data from 2010 has been estimated with informative prior information obtained as posterior information from model for data from 2009. Due to the change in the definition of levels for variable education in 2010 as compared to 2009, in the third model non-informative prior distributions have been used for this variable. The second model may be used as a reference model for the third one.

Before undertaking any inference from posterior distribution the convergence of generated Markov chains has been verified by the Geweke test [9], (see table 2). For all the parameters of investigated models no indication has been found that the Markov chains have not converged at any level of significance. The convergence of Markov chains has been confirmed by other tests and trace plots.

Table 2. The Geweke test

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1</th>
<th></th>
<th>Model 2</th>
<th></th>
<th>Model 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$z$</td>
<td>Pr &gt;</td>
<td>$z$</td>
<td>Pr &gt;</td>
<td>$z$</td>
<td>Pr &gt;</td>
</tr>
<tr>
<td>intercept</td>
<td>-1.5608</td>
<td>0.1186</td>
<td>0.1383</td>
<td>0.8900</td>
<td>0.9359</td>
<td>0.3493</td>
</tr>
<tr>
<td>sex1</td>
<td>0.5340</td>
<td>0.5933</td>
<td>0.0702</td>
<td>0.9440</td>
<td>-0.9284</td>
<td>0.3532</td>
</tr>
<tr>
<td>marital_status1</td>
<td>0.7778</td>
<td>0.4367</td>
<td>0.3511</td>
<td>0.7255</td>
<td>0.3913</td>
<td>0.6955</td>
</tr>
<tr>
<td>education1</td>
<td>-0.0130</td>
<td>0.9897</td>
<td>-0.3186</td>
<td>0.7500</td>
<td>-0.3132</td>
<td>0.7541</td>
</tr>
<tr>
<td>education2</td>
<td>-0.2544</td>
<td>0.7992</td>
<td>-0.1022</td>
<td>0.9186</td>
<td>1.0149</td>
<td>0.3102</td>
</tr>
<tr>
<td>education3</td>
<td>-0.0360</td>
<td>0.9713</td>
<td>-0.9752</td>
<td>0.3295</td>
<td>0.2080</td>
<td>0.8352</td>
</tr>
<tr>
<td>education4</td>
<td>0.1872</td>
<td>0.8515</td>
<td>-1.1732</td>
<td>0.2407</td>
<td>-0.5251</td>
<td>0.5995</td>
</tr>
<tr>
<td>education5</td>
<td>0.7346</td>
<td>0.4626</td>
<td>-1.4060</td>
<td>0.1597</td>
<td>-0.3220</td>
<td>0.7475</td>
</tr>
<tr>
<td>age_group1</td>
<td>1.2325</td>
<td>0.2178</td>
<td>0.5944</td>
<td>0.5523</td>
<td>0.6107</td>
<td>0.5414</td>
</tr>
<tr>
<td>region1</td>
<td>1.5252</td>
<td>0.1272</td>
<td>0.1809</td>
<td>0.8565</td>
<td>-0.7499</td>
<td>0.4533</td>
</tr>
<tr>
<td>region2</td>
<td>0.7866</td>
<td>0.4315</td>
<td>0.3710</td>
<td>0.7106</td>
<td>-1.3048</td>
<td>0.1920</td>
</tr>
<tr>
<td>region3</td>
<td>0.9525</td>
<td>0.3409</td>
<td>0.0646</td>
<td>0.9485</td>
<td>-0.1916</td>
<td>0.8480</td>
</tr>
<tr>
<td>region4</td>
<td>1.1261</td>
<td>0.2601</td>
<td>0.0439</td>
<td>0.9650</td>
<td>-0.4046</td>
<td>0.6858</td>
</tr>
<tr>
<td>region5</td>
<td>1.2234</td>
<td>0.2212</td>
<td>-0.2813</td>
<td>0.7785</td>
<td>0.4374</td>
<td>0.6619</td>
</tr>
<tr>
<td>place_of_living1</td>
<td>1.6682</td>
<td>0.0953</td>
<td>-0.0351</td>
<td>0.9720</td>
<td>1.4024</td>
<td>0.1608</td>
</tr>
<tr>
<td>place_of_living2</td>
<td>1.8743</td>
<td>0.0609</td>
<td>0.8223</td>
<td>0.4109</td>
<td>-1.3640</td>
<td>0.1726</td>
</tr>
</tbody>
</table>

In Table 3, the 95% highest posterior density regions (HPD) have been given for all models. These highest posterior density regions may be interpreted as: there is a 95% chance that the unknown parameters are in these regions. For any parameters (for example: marital status) large region differences for the first and the second models have been obtained. This may indicate that information from only one sample is insufficient to investigate a given occurrence. As investigated models are based on data from two successive years, obtained results should be similar, because no significant changes in the labour market were observed in this period. Including informative prior from the previous year in the analysis has improved the accuracy of estimation. The lower range of credible regions has been obtained in the model with informative prior distributions (Model 3) compared to model with non-informative prior distributions (Model 2), (see Table 4).

Table 3. The 95% highest posterior density regions (HPD)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>2.1541</td>
<td>2.1814</td>
<td>1.5498</td>
</tr>
<tr>
<td>sex1</td>
<td>-0.4235</td>
<td>-0.4068</td>
<td>-0.5783</td>
</tr>
<tr>
<td>marital_status1</td>
<td>-0.0954</td>
<td>-0.0782</td>
<td>-0.3947</td>
</tr>
<tr>
<td>education1</td>
<td>2.4528</td>
<td>2.4863</td>
<td>2.2231</td>
</tr>
<tr>
<td>education2</td>
<td>0.5851</td>
<td>0.6226</td>
<td>2.5278</td>
</tr>
<tr>
<td>education3</td>
<td>1.4139</td>
<td>1.4352</td>
<td>1.3003</td>
</tr>
<tr>
<td>education4</td>
<td>0.8189</td>
<td>0.8479</td>
<td>0.9101</td>
</tr>
<tr>
<td>education5</td>
<td>0.8180</td>
<td>0.8358</td>
<td>0.9436</td>
</tr>
<tr>
<td>age_group1</td>
<td>0.7404</td>
<td>0.7601</td>
<td>0.3802</td>
</tr>
<tr>
<td>region1</td>
<td>-0.0972</td>
<td>-0.0747</td>
<td>0.1604</td>
</tr>
<tr>
<td>region2</td>
<td>0.1592</td>
<td>0.1838</td>
<td>0.3281</td>
</tr>
<tr>
<td>region3</td>
<td>-0.2341</td>
<td>-0.2092</td>
<td>-0.1092</td>
</tr>
<tr>
<td>region4</td>
<td>0.0429</td>
<td>0.0675</td>
<td>0.4599</td>
</tr>
<tr>
<td>region5</td>
<td>-0.1407</td>
<td>-0.1158</td>
<td>0.0963</td>
</tr>
<tr>
<td>place_of_living1</td>
<td>-0.1644</td>
<td>-0.1457</td>
<td>0.1847</td>
</tr>
<tr>
<td>place_of_living2</td>
<td>-0.3791</td>
<td>-0.3618</td>
<td>-0.0895</td>
</tr>
</tbody>
</table>

Source: own analysis of the data "Household budgets" 2009 and 2010

The measure of simulation accuracy is Monte Carlo standard error (MCSE). The values of Monte Carlo standard errors for the analysed model are presented in Table 5. These results explicitly indicate that the best estimations have been obtained for the third model with informative prior distribution.
### Table 4. The length of the highest posterior density regions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>0.0213</td>
<td>0.0075</td>
</tr>
<tr>
<td>sex1</td>
<td>0.0135</td>
<td>0.0097</td>
</tr>
<tr>
<td>marital_status1</td>
<td>0.0132</td>
<td>0.0012</td>
</tr>
<tr>
<td>education1</td>
<td>0.0268</td>
<td>0.0247</td>
</tr>
<tr>
<td>education 2</td>
<td>0.0472</td>
<td>0.0458</td>
</tr>
<tr>
<td>education 3</td>
<td>0.0169</td>
<td>0.0149</td>
</tr>
<tr>
<td>education 4</td>
<td>0.0232</td>
<td>0.0217</td>
</tr>
<tr>
<td>education 5</td>
<td>0.0146</td>
<td>0.0130</td>
</tr>
<tr>
<td>age_group1</td>
<td>0.0142</td>
<td>0.0026</td>
</tr>
<tr>
<td>region1</td>
<td>0.0183</td>
<td>0.0001</td>
</tr>
<tr>
<td>region2</td>
<td>0.0191</td>
<td>0.0013</td>
</tr>
<tr>
<td>region3</td>
<td>0.0191</td>
<td>0.0001</td>
</tr>
<tr>
<td>region4</td>
<td>0.0203</td>
<td>0.0117</td>
</tr>
<tr>
<td>region5</td>
<td>0.0203</td>
<td>0.0112</td>
</tr>
<tr>
<td>place_of_living1</td>
<td>0.0150</td>
<td>0.0002</td>
</tr>
<tr>
<td>place_of_living2</td>
<td>0.0137</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

*Source: own analysis of the data "Household budgets" 2009 and 2010*

### Table 5. The values of Monte Carlo standard errors

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>0.00031</td>
<td>0.0437</td>
<td>0.00023</td>
</tr>
<tr>
<td>sex1</td>
<td>0.00013</td>
<td>0.0298</td>
<td>0.00010</td>
</tr>
<tr>
<td>marital_status1</td>
<td>0.00008</td>
<td>0.0187</td>
<td>0.00006</td>
</tr>
<tr>
<td>education1</td>
<td>0.00014</td>
<td>0.0157</td>
<td>0.00012</td>
</tr>
<tr>
<td>education 2</td>
<td>0.00015</td>
<td>0.0155</td>
<td>0.00014</td>
</tr>
<tr>
<td>education 3</td>
<td>0.00012</td>
<td>0.0212</td>
<td>0.00010</td>
</tr>
<tr>
<td>education 4</td>
<td>0.00014</td>
<td>0.0181</td>
<td>0.00011</td>
</tr>
<tr>
<td>education 5</td>
<td>0.00011</td>
<td>0.0238</td>
<td>0.00009</td>
</tr>
<tr>
<td>age_group1</td>
<td>0.00006</td>
<td>0.0120</td>
<td>0.00005</td>
</tr>
<tr>
<td>region1</td>
<td>0.00017</td>
<td>0.0296</td>
<td>0.00012</td>
</tr>
<tr>
<td>region2</td>
<td>0.00016</td>
<td>0.0261</td>
<td>0.00012</td>
</tr>
<tr>
<td>region3</td>
<td>0.00017</td>
<td>0.0270</td>
<td>0.00012</td>
</tr>
<tr>
<td>region4</td>
<td>0.00017</td>
<td>0.0274</td>
<td>0.00013</td>
</tr>
<tr>
<td>region5</td>
<td>0.00016</td>
<td>0.0258</td>
<td>0.00013</td>
</tr>
<tr>
<td>place_of_living1</td>
<td>0.00011</td>
<td>0.0223</td>
<td>0.00008</td>
</tr>
<tr>
<td>place_of_living2</td>
<td>0.00010</td>
<td>0.0226</td>
<td>0.00008</td>
</tr>
</tbody>
</table>

*Source: own analysis of the data "Household budgets" 2009 and 2010*
The interval estimation does not give information which values from credible intervals are the most probable, therefore empirical results have been supplemented by the point estimation values (see Table 6). In this paper, only the results for model 3 with informative prior distributions have been presented. Based on the highest probability density interval [3], all variables are statistically significant. The odds ratio has been calculated as well for a more detailed interpretation of the obtained results.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>sex1</td>
<td>-0.5046</td>
<td>0.00245</td>
<td>0.604</td>
</tr>
<tr>
<td>marital_status1</td>
<td>-0.2592</td>
<td>0.00041</td>
<td>0.772</td>
</tr>
<tr>
<td>education1</td>
<td>2.2374</td>
<td>0.00638</td>
<td>9.369</td>
</tr>
<tr>
<td>education 2</td>
<td>2.5794</td>
<td>0.01170</td>
<td>13.189</td>
</tr>
<tr>
<td>education 3</td>
<td>1.3455</td>
<td>0.00381</td>
<td>3.840</td>
</tr>
<tr>
<td>education 4</td>
<td>0.9494</td>
<td>0.00556</td>
<td>2.584</td>
</tr>
<tr>
<td>education 5</td>
<td>0.9859</td>
<td>0.00333</td>
<td>2.680</td>
</tr>
<tr>
<td>age_group1</td>
<td>0.5062</td>
<td>0.00084</td>
<td>1.659</td>
</tr>
<tr>
<td>region1</td>
<td>0.3734</td>
<td>0.00004</td>
<td>1.453</td>
</tr>
<tr>
<td>region2</td>
<td>0.3332</td>
<td>0.00043</td>
<td>1.395</td>
</tr>
<tr>
<td>region3</td>
<td>0.1502</td>
<td>0.00004</td>
<td>1.162</td>
</tr>
<tr>
<td>region4</td>
<td>0.4132</td>
<td>0.00313</td>
<td>1.512</td>
</tr>
<tr>
<td>region5</td>
<td>0.0931</td>
<td>0.00306</td>
<td>1.098</td>
</tr>
<tr>
<td>place_of_living1</td>
<td>0.1564</td>
<td>0.00007</td>
<td>1.169</td>
</tr>
<tr>
<td>place_of_living2</td>
<td>-0.1403</td>
<td>0.00017</td>
<td>0.869</td>
</tr>
</tbody>
</table>

Source: own analysis of the data "Household budgets" 2010

The obtained values indicate that men have 39% less chance of having a job than women. For unmarried people the chance of having a job is about 23% less than for married people. The individuals who have an education level higher than primary, are more likely to have a job; 168% more likely for basic vocational, 158% more likely for secondary general, 284% more likely for secondary professional. Moreover, it is over thirteen times more probable for individuals who have post-secondary education to have a job than for people with the lowest education level. For people with higher education it is over nine times more probable.

People who live in other region of Poland than north have more chance of having a job. Moreover, people who live in the central region of Poland have the biggest chance of having a job as compared to people who live in north. As for the place of living, people who live in big cities are 16.9% more likely to have a job than people who live in the country and people who live in small cities are 13.1%
less likely compared to people who live in the country. People aged 60 and older have 65.9% more chance of having a job than people who are under 60.

5. Conclusion

In this paper Bayesian confidence intervals have been investigated in the context of logistic regression model. This approach entails significantly larger computational costs than classical methods due to a higher model complexity. The models have been estimated using Markov chain Monte Carlo methods with Gibbs sampling. In this work, benefits arising from the use of Bayesian approach to modelling, especially when confidence intervals are determined, have been shown.

The primary advantage of the proposed approach is the ability to use out of the sample knowledge in the modelling process. This is particularly useful when modelling is performed on a regular basis as in the case of "Household budgets" research. The use of information from the preceding year yielded better parameter estimation. In particular, compared to the use of non-informative prior distributions, lower values of Monte Carlo standard errors have been attained.

Bayesian confidence regions yielded information on the range of change of estimated parameters with probability of 0.95. Moreover, this result was obtained based on this particular sample, while using a priori information from the previous survey. The use of informative prior distribution resulted in the significant reduction of highest posterior density region range as compared to the model using non-informative prior distribution.

Based on the estimation made, the feature that influences most the employment status of the persons 55 years old and older is education status. The results for the sex variable are different from the results of other studies, conducted for the entire population, for example from BAEL research results [18].

REFERENCES


EQUIPMENT LEVEL OF POLISH HOUSEHOLDS IN INFORMATION TECHNOLOGY

MAŁGORZATA GRZYWIŃSKA-RAPCA

Department of Quantitative Methods,
University of Warmia and Mazury in Olsztyn (UWM)

The problem of the use of information technology can be analyzed in terms of aspirations, goals and values, but also in terms of the region’s population to participate actively in the development of the information society. Statistical studies show different levels of access to information technologies. To actively participate in the surrounding reality beyond the skills of the users of communication technology is an important access to modern information technology. The aim of this article is to assess the level of the use of modern technology in households.

Keywords: households, Internet users, levels of access to information technologies

1. Introduction

Modern technology offers users a variety of specialized functional communication tools with computer networks, especially the Internet. Effective network access is a prerequisite for the implementation of suit your preference [3, 5]. Modern technology is a factor in determining the direction of the transformation of contemporary developed societies. A study conducted by TNS Digital Life shows that in Poland the average Internet user spends online an average of 18 hours and 25 minutes per week. It’s almost 80 hours per month and almost 1,000 hours a year, which translates into 40 days of continuous residence before the monitor [11]. Information needs are so important in the hierarchy of needs of a modern household [5]. The level of their satisfaction affects the
development of the individual units, and the operation of the entire household. It is
determined weave a variety of conditions and factors that influence often difficult
to identify and explore. The aim of the study is to assess the level of use of the
equipment and Internet access to Polish households.

2. The level of access of EU households

The dynamic development of information and communication technologies
since 2003 for analysis of access and use of modern technology has been
developed digital access index (DAI) based on four factors affecting the
development of the information society of the country [12]. These factors included
infrastructure, affordability, knowledge and skills, and the actual use of ICT -
communication. Among the 181 countries of the EU countries occupy top positions
in the ranking. Regardless of the type and classification of factors influencing the
development of the information society is the most important equipment of
households with computers and Internet access. The level of equipment of
households with a computer with Internet access is shown in Figure 1.

![Figure 1. Households with broadband Internet access in selected European countries](image)

In 2013, the highest percentage of households with broadband Internet access
were recorded in Ireland, Denmark and Finland. At least in the high-speed Internet
access are provided with households Romania. Poland despite the increase in the
level of equipment over the years 2004-2013 in the ranking it gained slightly below
the European average.
3. Polish household equipment in hardware

Development of telecommunications, including the proliferation of wireless telephony Polish consumers will be from year to year are increasingly used the Internet not only as a source of information, but also as a tool to facilitate action in everyday life [1]. The results of a study conducted by the Central Statistical Office on household equipment show an improvement in the level of equipment in hardware [4]. The most important factor influencing the access to the global network is to provide households with a computer. Percentage of households with a computer has steadily risen in recent years. While in 2009, 66.1% of households were equipped with computers, in the next year already 69%. The lowest increase in the number of households with a personal computer appeared in Eastern Poland (an increase of 7 percentage points in 2013 with respect to 2009), while the highest in central Poland - 9 pp (Information Society in Poland, 2013).

Table 1. Equipment Polish households with a personal computer in 2004-2013

<table>
<thead>
<tr>
<th>specification</th>
<th>Total personal computer [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Łódzkie</td>
<td>30.4</td>
</tr>
<tr>
<td>Mazowieckie</td>
<td>35.8</td>
</tr>
<tr>
<td>Małopolskie</td>
<td>36.9</td>
</tr>
<tr>
<td>Śląskie</td>
<td>34.1</td>
</tr>
<tr>
<td>Lubelskie</td>
<td>31.2</td>
</tr>
<tr>
<td>Podkarpackie</td>
<td>30.5</td>
</tr>
<tr>
<td>Podlańskie</td>
<td>32.4</td>
</tr>
<tr>
<td>Świętokrzyskie</td>
<td>26.3</td>
</tr>
<tr>
<td>Lubuskie</td>
<td>29.8</td>
</tr>
<tr>
<td>Wielkopolskie</td>
<td>34.0</td>
</tr>
<tr>
<td>Zachodniopomorskie</td>
<td>32.0</td>
</tr>
<tr>
<td>Dolnośląskie</td>
<td>31.4</td>
</tr>
<tr>
<td>Opolskie</td>
<td>31.6</td>
</tr>
<tr>
<td>Kujawsko-Pomorskie</td>
<td>30.8</td>
</tr>
<tr>
<td>Pomorskie</td>
<td>36.5</td>
</tr>
<tr>
<td>Warmińsko-Mazurskie</td>
<td>27.2</td>
</tr>
</tbody>
</table>

Source: own preparation on basic [8]
Household equipment varies depending on the number of children in the family, or, depending on income. An important factor differentiating the level of equipment of households is also a place of residence (Table 1).

The largest increase in households with a personal computer in relation to 2004 was observed in świętokrzyskie. This change was at 42.4 p.p. The smallest change in the equipment of households with a personal computer was reported in podlaskie (32 pp). It should be noted that in 2013 the average level of equipment in a personal computer in Poland was 71.5% (Fig. 2).

![Figure 2. Level of personal computer equipment in Poland in 2013](image)

Province, where the level of personal computer equipment in 2013 was higher than the national average are: mazowieckie, małopolskie, śląskie, wielkopolskie and pomorskie. Pomorskie the percentage of households with a personal computer was a 5.8 pp higher than the national average and amounted to 77.3%. At the lowest level of personal computer equipment was in podlaskie - 64.4%.

4. Internet in households

One of the mainstays of various sectors of the economy is a better use of information technology. The methodology research in the European Union, as a household with Internet access takes the farm and the possibility of having the computer communications [10]. Among households equipped with a means of
access to the Internet is still a large group without Internet connectivity. Table 2 shows the level of equipment with a computer with Internet access in 2004-2013.

Table 2. The level of equipment of Polish households with a computer with internet access in 2004-2013

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Łódzkie</td>
<td>14.4</td>
<td>18.4</td>
<td>25.4</td>
<td>32.6</td>
<td>40.6</td>
<td>47.0</td>
<td>52.3</td>
<td>54.9</td>
<td>59.2</td>
<td>62.8</td>
</tr>
<tr>
<td>Mazowieckie</td>
<td>20.2</td>
<td>26.9</td>
<td>32.5</td>
<td>40.5</td>
<td>49.4</td>
<td>56.5</td>
<td>64.2</td>
<td>67.5</td>
<td>69.2</td>
<td>73.0</td>
</tr>
<tr>
<td>Małopolskie</td>
<td>20.0</td>
<td>27.7</td>
<td>32.5</td>
<td>41.8</td>
<td>51.0</td>
<td>57.4</td>
<td>63.6</td>
<td>64.2</td>
<td>67.2</td>
<td>72.7</td>
</tr>
<tr>
<td>Śląskie</td>
<td>18.3</td>
<td>27.3</td>
<td>33.9</td>
<td>42.2</td>
<td>51.3</td>
<td>57.6</td>
<td>64.1</td>
<td>65.2</td>
<td>67.9</td>
<td>72.9</td>
</tr>
<tr>
<td>Lubelskie</td>
<td>14.6</td>
<td>16.4</td>
<td>19.4</td>
<td>28.1</td>
<td>36.4</td>
<td>46.0</td>
<td>50.6</td>
<td>55.4</td>
<td>59.5</td>
<td>65.4</td>
</tr>
<tr>
<td>Podkarpackie</td>
<td>15.8</td>
<td>22.0</td>
<td>24.8</td>
<td>30.9</td>
<td>41.9</td>
<td>50.4</td>
<td>59.5</td>
<td>62.2</td>
<td>62.6</td>
<td>67.6</td>
</tr>
<tr>
<td>Podlaskie</td>
<td>17.4</td>
<td>19.7</td>
<td>28.0</td>
<td>33.9</td>
<td>44.1</td>
<td>51.5</td>
<td>52.6</td>
<td>57.5</td>
<td>59.4</td>
<td>61.0</td>
</tr>
<tr>
<td>Świętokrzyskie</td>
<td>11.0</td>
<td>13.5</td>
<td>18.2</td>
<td>22.2</td>
<td>32.5</td>
<td>43.3</td>
<td>52.2</td>
<td>52.5</td>
<td>56.7</td>
<td>64.3</td>
</tr>
<tr>
<td>Lubuskie</td>
<td>13.1</td>
<td>24.5</td>
<td>31.1</td>
<td>34.7</td>
<td>45.8</td>
<td>55.9</td>
<td>60.5</td>
<td>59.9</td>
<td>59.7</td>
<td>65.9</td>
</tr>
<tr>
<td>Wielkopolskie</td>
<td>14.1</td>
<td>20.4</td>
<td>26.7</td>
<td>36.5</td>
<td>42.9</td>
<td>51.0</td>
<td>59.4</td>
<td>64.3</td>
<td>68.5</td>
<td>71.7</td>
</tr>
<tr>
<td>Zachodniopomorskie</td>
<td>16.2</td>
<td>20.3</td>
<td>27.6</td>
<td>35.6</td>
<td>46.4</td>
<td>55.3</td>
<td>57.4</td>
<td>61.0</td>
<td>64.0</td>
<td>65.5</td>
</tr>
<tr>
<td>Dolnośląskie</td>
<td>17.1</td>
<td>22.7</td>
<td>28.4</td>
<td>36.2</td>
<td>46.4</td>
<td>53.6</td>
<td>59.3</td>
<td>60.9</td>
<td>62.6</td>
<td>66.0</td>
</tr>
<tr>
<td>Opolskie</td>
<td>13.1</td>
<td>19.3</td>
<td>23.9</td>
<td>34.3</td>
<td>44.1</td>
<td>52.5</td>
<td>57.3</td>
<td>63.4</td>
<td>66.7</td>
<td>66.5</td>
</tr>
<tr>
<td>Kujawsko-Pomorskie</td>
<td>14.9</td>
<td>16.8</td>
<td>20.8</td>
<td>33.6</td>
<td>43.1</td>
<td>53.7</td>
<td>58.6</td>
<td>59.5</td>
<td>59.2</td>
<td>64.5</td>
</tr>
<tr>
<td>Pomorskie</td>
<td>21.8</td>
<td>24.8</td>
<td>37.3</td>
<td>47.0</td>
<td>53.0</td>
<td>58.0</td>
<td>64.2</td>
<td>69.4</td>
<td>71.0</td>
<td>76.1</td>
</tr>
<tr>
<td>Warmińsko-Mazurskie</td>
<td>13.9</td>
<td>17.8</td>
<td>22.1</td>
<td>28.7</td>
<td>39.0</td>
<td>48.8</td>
<td>55.8</td>
<td>58.1</td>
<td>58.2</td>
<td>61.5</td>
</tr>
</tbody>
</table>

Source: own preparation on basic [9]

In Poland, in 2013 the percentage of households with access to the Internet, in which at least one person has a global network was 71.9%. The highest increase in households equipped with a computer with Internet access during the Polish EU membership was observed in Wielkopolska by 57.6 percentage points (from 14.1% in 2004 to 71.7% in 2013). Despite various efforts to bring the level of household equipment Polish Internet access varies between regions (Fig. 2).

In 2013, 68.8% of households had access to the Internet. At the slightest possibility of using the network to residents Podlaskie (61%), warmińsko-mazurskie (61.5%) and łódzkie (62.8%).

Because of the speed and quality of information flow every Internet user seeks to obtain broadband. This is a link increasingly being used by users, but for various reasons, there are still many households do not have such access to the Internet. The level of equipment in the Polish broadband households in 2013 is shown in Figure 3.
Figure 3. The level of equipment of households with a computer with Internet access

Figure 4. The level of equipment in the Polish broadband households in 2013

In Poland in 2013, households have broadband access to the Internet accounted for 55.4%. The highest percentage of users were recorded in Podkarpackie province, where 67.6% of households have broadband Internet.
access. Households from świętokrzyskie are but the least equipped with broadband access. Only 35.2% of individuals has the opportunity to reach the network.

5. The objectives use of the Internet

Good quality-Speed Internet allows users to better meet your needs. Despite the increasing range of products and services offered through the Internet, Polish users for several years in a study conducted by the Central Statistical Office show as the most frequently performed operation in the network to send and receive electronic mail (Fig. 4).

<table>
<thead>
<tr>
<th>Activity</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sending and receiving e-mail</td>
<td>52.9%</td>
</tr>
<tr>
<td>Participation in chat rooms, discussion groups or forums, instant messaging, creating and reading blogs and the use of social networking sites</td>
<td>41.1%</td>
</tr>
<tr>
<td>Calling over the Internet, video conferencing</td>
<td>27.5%</td>
</tr>
<tr>
<td>Searching for information about goods or services</td>
<td>50.1%</td>
</tr>
<tr>
<td>Using the services of services related to travel and accommodation</td>
<td>13.8%</td>
</tr>
<tr>
<td>Listening to Internet radio</td>
<td>19%</td>
</tr>
<tr>
<td>Downloading software</td>
<td>14.9%</td>
</tr>
<tr>
<td>Downloading files from games, music or video</td>
<td>27.6%</td>
</tr>
<tr>
<td>Reading online, download files from newspapers or magazines</td>
<td>47.2%</td>
</tr>
<tr>
<td>Banking services</td>
<td>32.6%</td>
</tr>
<tr>
<td>Finding a job, sending offers for employment</td>
<td>11.6%</td>
</tr>
<tr>
<td>Sales of goods or services</td>
<td>11.6%</td>
</tr>
</tbody>
</table>

Figure 5. Most frequent activities on the Internet
Second order, the use of e-mail, Polish Internet users find information exchanged target about goods and services (50.1%), while it was not clear to the purchase. A large group of surfers uses modern technology to read online and download magazines or newspapers (47.2%).

6. Conclusions

In developing the information society household equipment can be reflected in the emergence of new needs of the applications of ICT. The availability of modern technology in households is a factor contributing to the new behavior and changes in the mentality of household members, the organization purchases and consumption. The availability of modern technology gives you opportunities, but we should not forget the limitations of household income. It should be remembered, as confirmed by the data in Table 2 that there are significant differences in the equipment of information and communication technologies by households, depending on their location. Generally, the state of the IT hardware equipment is a phenomenon as an expression of the most positive and getting better meet consumer needs of Polish households in this area. Equipped with modern technology contributes to changes in many aspects of life. Among them, the great hope for the future, for example, a change in the way the trade. E-commerce allows you to make a completely digitized business processes using a global network the Internet. The development of e-services and e-commerce therefore also strongly developed in the European Union [2]. E-services is to be the way to improving the quality of life of societies and increasingly better orientation to satisfy all the needs of consumers [6]. With e-commerce, the market increases the number of available information, increasing the speed of circulation. Transactions are processed on the global market, and are therefore not dependent in any way from the border states [7].

REFERENCES


FUNCTIONAL ANALYSIS OF INFORMATION SYSTEMS SUPPORTING THE PROCESS OF FACILITY MANAGEMENT

MONIKA JARANOWSKA

Department of Management, Warsaw University of Technology

Application of supporting management processes through the use of information systems is now becoming one of the basic requirements of market competition. Professional distributed systems not only facilitate collaboration with customers but also reduce operating costs. The aim of this analysis is to present the functionalities that may characterize management system, indicating their relevance to property managers. Market overview confront the expected functionality with the practical implementation of the identified solutions for the most popular systems for universal use. The summary of the significance of using CAFM software will be the review of selected commercial and municipal projects.

Keywords: IT systems, property management, support the management, functional analysis, IT audit

1. The essence of property management systems

Power management with the use of information systems is a mandatory approach. Practically every mainstream enterprise or organization shall apply such tools, using universal solutions or projects copyright.

For the purposes of the Land Administration Act dated August 21, 1997 year, the property management is to take all decisions and make all activities aimed at maintaining real estate in not deteriorated condition in accordance with its intended purpose, as well as a reasonable investment in the property. Permanent part of the practical work property manager is not only taking care of the maintenance of
facilities supplies function at the highest level, but also repetitive administrative tasks – financial and accounting controls and, above all strategic activities [3]. That are conducive information systems, which apart support these actions can be characterized by elements of analytical and decision-making based on the model of business intelligence. This eliminates not only the number of specialists required for operation of the board, but also reduces the risk of error generating activities of employees in the area of repetitive tasks.

Problem perfection system is based on the ability of the project to map business processes. Real estate management support systems are among the PMS - property management systems. The structure is based on relationship the client - server, but now increasingly popular solutions are those based on the cloud (as a service). PMS solutions are most often used by local governments, commercial property managers, hotel industry, largely residential and developers community [3].

2. Requirements for CAFM

Valid system should ensure the flow between the modules technical, administrative, financial and analytical in a manner transparent to the user and gives the greatest spectrum of possibilities. Users should only monitor the correctness of the results and generate conclusions about the current state of the property. Extension of the system should also be communication with the customer, allowing for fault reporting or information about problems. This reduces significantly the number of phone calls and absorption of employees.

Currently, the main determinant of choice for property management system is the price factor. For projects with CAFM (Computer – Aided Facility Management) group are distinguished: the price of the license and price of the implementation. The license price is quite disputable criterion - the value of the license should be a fraction of the cost of the management, but it is not a priority over the choice of system functionality and technological advancement. In the case of implementation, it is a factor that should be minimized, because the costs of training are often overestimated, while too low a level of quality offered.

It can be concluded that the primary functions CAFM systems are the absolute minimum requirements for software vendors. There are modules on real estate records, premises and tenants, including basic data on the managed assets and clients. Based on the database module operates clearing rents, which in addition to the classic records of claims and payments should work with the creation of any billing algorithms (keys), enabling the sharing of common costs to tenants.

With rents module should be coupled module utility billing and accounting counters. Regardless of the method for data extraction technology, costs should be
allocated to the tenant and settled in liabilities. Should also take into account the different forms of grants, consistent with the provisions of nationwide and local (housing allowances) and interest notes.

Modules records and rents should be linked to electronic banking module - home banking. Automated record keeping and accounting transfers and their automatic execution reduces the risk of human error. A convenient aspect would be also link the described module with the accountant module [12].

In terms of financial - accounting, the expected functionality is fully compliant with the current Accounting Act. In this area, the system should recognize the costs and revenues, including created a chart of accounts and making support for the tax.

Rental module is complemented by vindication module. It has not only analyse customer commitment to the administrator, but also include a legal aspect to recovery before the court, judicial and debt collectors.

In the area there is a wide economic additional functionality, i.e. invoicing module, recording cash services (if such is necessary in the company) or payroll - referring to the cost of the administrator, incurred in connection with the operation real estate [2].

Almost every business and property management functionality requires supervision and control. These elements ensure success and ability to raise additional funds for investment and improve their own situation in a competitive market. CAFM system should have a highly flexible and customizable module for controlling, supported by elements of the statistical analysis. The most convenient form should be a model of business intelligence, data warehousing and OLAP (OnLine Analytical Processing) service. Basic information, such as vacancy rates and average rents in the area are not already sufficient, and predefined analytical models are evolving. Management is using cartogram facilitating the addition of visual presentation.

In addition to the functionality of the system and reporting modules (which will undoubtedly facilitate the functioning), it becomes necessary to export the data to external systems (even so popular Excel spreadsheet). Of course, export format should meet the requirement of universality and be easily transferable between systems.

Administrative aspect mainly includes contact with clients, customers and internal circulation of documents. Internal module streamlines document workflow actions, accelerates the process of decision-making, and what is most important functions pro ecologically. Of course, it is necessary to extend the functionality of the legal unit, including management agreements and obligations, tenders, adopted resolutions and planned repairs, renovations or upgrades.
Complemented document circulation is the system of communication with partners and customers, both the web site and the internal email system. Communication is understood not only in the form of active and passive but with access to the assigned groups of information and the ability to change certain parameters relating to the lease, without the involvement of the employee. This reduces not only the costs but the general management and minimizes the time of customers, which is very appreciated aspect.

Typically administrative functionality linked with automated information generated by the system and the need to provide specific facts audience is serial correspondence (and a record of correspondence). Its classic form can be extended to notification by email and messages. For each of the users of the system would be a convenient addition of a typical PIM (Personal Information Manager), with an alert function on significant dates, personalization and full adaptation to specific audiences.

Of course cannot forget about such a fundamental element of which is the technical records of objects and places of module failure and planned investments. The ability to handle CAD (Computer Aided Design) documentation and survey data in the form of GIS (Geographic Information System) digital mapping and storage of energy certificates is very convenient [2].

The technology which software is made of one of the most important criteria for choosing a solution for the company. The system should use the latest technology and information which will allow the company, at no additional cost, to have the current arrangements for an extended period of time. Equally important is the ability to easily find the implementation of new modules and upgrades, which will increase the level of safety for the services offered. Currently, the most popular and most practical theory subscription fee for the use of the software in the cloud. A company that provides software provides also a high level of security and ensure quality of service, but the long-term costs may be higher here than those incurred in the classical licenses.

Referring to the technological element and extend the capabilities of specific additives interesting is the recognition of its own programming language in the system. It will facilitate the parameter setting via scripts written by our own IT team, without having to wait for producer responses for the implementation of new features.

The solution to which a few years ago it seemed literary fiction is the use of telemetry and remote sensing. Equipment associated with the system will not be able to read utility meters, checking the technical condition of buildings and detect failures, but above all, give a chance for automated responses and corrective actions to protect the object with the users. At stake is not only a safety issue but
also the convenience of customers who are not cyclically engaged i.e. the meter readings.

Very important, but often downplayed element is the user interface. Its ergonomics and simplification will help to reduce the cost of training staff and decrease reluctance to new solutions. In addition, customers and partners with more gladly accept the possibility of using the system to carry out certain tasks or communication itself. Important here is the flexibility and freedom to adapt to individual user preferences.

Relationships and CAFM systems requirements is presented in Figure 1, creating a kind of diagram to the expectations of client software provider.

![Figure 1. Requirements and relationships for CAFM software. Source: Gravity Group](image)

Unquestionable support but a necessity arising from the popularization of management systems becomes an independent oversight of the software, regardless of whether it relates to real estate or any other sector. Ensuring compliance objectives of the company or other institution and the efforts of IT to support the
management and diagnosis of the current state of information technology and matching business activities, is auditing information systems. Currently the most popular institution acting in this area is ISACA (Information Systems Audit and Control Association). It allows the company to ensure the quality and effectiveness of the delivered system, increasing confidence in the compatibility of the requirements to the system at its actual functionality. Institutions and companies that decide to implement this solution or copyright, or universal, should conduct an audit of compliance in the areas of design, application and security [4].

3. A review of selected universal solutions

One of the most comprehensive property management systems is package Granit, that is addressed to a wide range of large and very large customers. It offers all the functionality associated with running a business management through classic modules, such as finance and accounting, community housing, rent, utilities, debt collection, repairs and investment, human resources and others. Innovation distinguished modules for generating reports reporting period of lower system load using the data warehouse to facilitate payment transactions contact the bank, electronic document circulation, collector of media including billing management (working with PocketPC devices) and two very powerful: a web portal for communication with tenants along with the history of the implementation of commitments and data relating to the premises and the module map, including support for digital maps [10].

Interesting and fairly well-known program is the Sizan offered by the Regional Centre for Informatics Ltd.. It is dedicated to local government units, HA – Facilities (Housing Association Facilities), cooperatives, communities, property management companies and companies in order to exploit for their own facilities. The system consists of five subsystems logically related dues, debt, financial - accounting, buildings, Records of the Communities and Buildings, and settlement. The system is not so complete and implementation may be subjected to a selected group of functions. Receivables subsystem includes all records and accounts to tenants, together with cash and bank modules allowing a comfortable operation and management of payments. Recovery subsystem includes legal elements (the case in the court, judicial and enforcement officers). Accountant finance module developed on the basis of the chart of accounts structure builds balance, accounts for the costs, hedge counterparties and attribution, and includes basic reporting tool. Registration Subsystem oversees energy certificates, specifies the technical parameters, failures, planning and filing repairs. Settlement module determines the number of units available, and it is associated with applications pending, also refers the situation in the real estate [9].
The solution based on the cloud Axxerion is the system, consisting of a very large number of modules and functionality. Property management module, contacts, contracts, proposals and employees complement each other allowing for access to the whole information about the managed properties. Technical support is key management module increases the level of security and flexibility of the project. Project management module facilitates investment operations manager and simplifies the financial analysis. The versatility allows you to use the software for commercial centers - with the function of entertainment events, offices and hotels. Rely on a cloud system increases safety and reduces the costs associated with maintaining their own servers.

TelkomBud program, is dedicated to commercial galleries and office objects. The executor of the solution is ProgPol company offering the product as highly customizable and able to link up with the current firmware (without migration). In addition to the functionality of a typical for this type of program, covers the requirements and deadlines, books of a building, a lease, operating agreement, or a data bank for land - related modules of the importance of investment activity [11].

Perseus WM program is quite rich functional but simple system, built on the basis of modules. The disadvantage is specific GUI (Graphical User Interface), that despite the readability of brings to mind previous systems. It contains the basic elements characteristic of these types of programs offering full compatibility with spreadsheets (Excel, Open Office). Additional modules of the system are: Home Banking with the possibility of cooperation with banks and automatic transfer posting, Group Office allows you to access the Internet, exposing module transfers and cash reports.

At the end of the review it is worth recalling the oldest application, which was the precursor to the development of CAFM systems in Poland. It is the ADA's Pixel, addressed to cooperatives and condominiums. This system is very limited in its design and technology, which is a form of data files with a common structure. Each module is a separate program ADA "rents" collecting charges for the use of the premises, "media" including data on collective and individual counters, "costs" relating to the operating costs of facilities, "cash", "program financially accountant" and "web "which financial module tenant files on the internet along with the history of payments. ADA package can also be expanded with additional modules contributions, DCM and cooperating number of separate programs. Unfortunately, quite out-dated structure makes it very difficult at present to use the tool, despite its versatility and flexibility.

All of these programs are characterized by a modular service. Only one supports only commercial real estate, while the other is characterized by high versatility. Still rare is CAD file and GIS map service. Summary comparison of functionality these systems is presented in the Table 1.
Table 1. Comparison of CAFM systems

<table>
<thead>
<tr>
<th>Developer</th>
<th>Granit</th>
<th>Sizan</th>
<th>Axserion</th>
<th>TelkomBud</th>
<th>Perseus</th>
<th>ADA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DomCon</td>
<td>WOI</td>
<td>Unima 2000</td>
<td>ProgPol</td>
<td>Perseus</td>
<td>Pixel</td>
</tr>
<tr>
<td>Module</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>rents</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>media</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>vindication</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>accountancy</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>renovations</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>real estate registry</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>staff</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>controlling</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>GIS map</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>CAD files</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>PIM module</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>internet portal</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>contract management</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Expendability</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLAP service</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target group</td>
<td>Housing commu nities, cooperatives, municipalities, companies</td>
<td>Housing communities, cooperatives, municipalities, companies</td>
<td>Housing communities, cooperatives, shopping malls, office buildings</td>
<td>Shopping malls, office buildings</td>
<td>Housing communities, cooperatives, accountancies</td>
<td>Housing communities, cooperatives</td>
</tr>
<tr>
<td>Additional functionality</td>
<td>Cooperation with the PocketPC</td>
<td>Key management system based on the cloud</td>
<td>The possibility of linking with the current system, fully compatible with MS Office</td>
<td>Full compatibility with MS Office, Home Banking</td>
<td>Modules as separate programs</td>
<td></td>
</tr>
</tbody>
</table>

Source: own preparation on the basis of data system developers
4. The popularity of CAFM software

Polish companies have introduced their own property management systems rather seldom, but they did not fully realize the tasks set by the authorities. One exception is the company Est-On and from 2011 Echo Investment Property Management, real estate management company which is one of the two largest Polish developers - Echo Investment. Solution based on a distributed architecture used there is very specific and original. It condenses the special characteristics of systems in the area of economic management - information, advisory, taking into account the technical aspect of the control subjects. The system complements the classical solution of the management software for control over the work of the technical installations which ensures uninterrupted operation of the object. So demanding specification is associated with the requirements of the client and the type of managed properties - office buildings and shopping centers, where a break in the normal functioning generate significant financial loss tenants. It is worth mentioning that the entity manages sixteen properties with a total area of 370 000 m2, performing services to 820 commercial tenants. Expenditure incurred for the construction of such an extensive system of multi-dimensional turn - both in terms of convenient and low-cost real estate management as well as from the point of view reputation building.

The rarity of such solutions in the enterprise may be the result of their innovative in the Polish market and the lack of original solutions. Universal software is often too inflexible and addressed to numerous but smaller customers, they are housing associations, community or logistics operators. Another reason for the lack of popularity may be a factor implementations that are virtually never-ending process, and quite expensive for the company. Group of factors that reduce the popularity of the software closes the lack of awareness on the part of the customer, in terms of improvements that the system can make to the functioning of the institution or company and the hostility of middle and lower administrative level, fearing for their positions [1].

PMS solution also uses a large number of local government units, mainly in the cities and municipalities. This simplifies the management of objects and part of the city's own municipal. Reduced costs and ease of operation of the offices of reaching out to the residents of difficult communication technology (elderly and disabled).

The first example of a fairly large involvement in the development of an integrated management system is real estate investment Krakow, who announced a tender for the realization of an information system for property management Małopolska and provincial units. The budget of the project, in this case amounted to 1.6 million PLN, being co-financed by the European Regional Development Fund (ERDF) under the Małopolska Regional Operational Programme for the years
2007 - 2013. The Company carries out the project was Granit FM. The system consisted of 11 modules, allowing you to create a coherent whole. One of the main elements of a resource management module, which is a database of administrative, technical, legal and surveying the possible cooperation with CAD data. The module supports the management of parcels, buildings and structures, using them for statutory purposes e.g. The deployment of personnel and equipment of their jobs, and also for auxiliary purposes, e.g. parking places, etc. Fixed asset management was covered in a separate module, allowing you to keep records and accounts assets in accordance with accounting principles. Other modules include records of people, routine procedures (collecting and supervising the implementation of the terms of cyclical and non-cyclical work on existing buildings) and counter module, which supports recording utility meters for each of the resources [5].

Radom is another city involving property management system for the administration of their own assets. Tender for the supply and implementation of property management information system technology client - server was announced in 2008. The distinguishing feature of the project were very detailed records of the Terms of Reference for the expectations of the system. The project budget amounted to 95 000 PLN, the executor was SoftHard company [6].

Experience gained in the implementation of the project for the Radom, SoftHard company used in the implementation of the corresponding solutions for Szczecin. The tender was realized in 2013, including the transfer of existing processes and functional subsystems used with the training of employees [7].

Łódź is the last described example of system implementation, whose size is quite large. In the management of cities and villages is 60 thousand. premises, 10 000 buildings, 6 000 real estate, and rental services are offered to 200 000 residents. The tender was announced in 2013, with a budget oscillating around 1 million PLN. The Company's planned project was Sacer [8].

5. Conclusion

In the paper are characterized the key elements of the functionality of the property management systems, aimed to support the work of property managers. Certain key requirements that favour contracting with respect to companies implementing software to support. It explains common property management systems, as well as provides examples of their use in real projects. The analysis clearly shows that the use of information systems has a positive effect on the efficiency and quality of property management. However, universal systems do not ensure the implementation of all the needs of users of the software. System to provide full support in all aspects of real estate management, must be created according to individual orders of the management unit, which however generates
significant costs. Therefore, few companies in Poland introduced their own property management systems, but uses them huge number of local government units.

REFERENCES


