DESIGN OF A LABORATORY INFORMATION SYSTEM FOR DATA PROCESSING AND EFFICIENCY EVALUATION

Vladimir GAŠPAR*, Rudolf ANDOGA**

*Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Letná 9, 042 00 Košice, Slovak Republic, tel. +421 55 602 4216, e-mail: vladimir.gaspars@tuke.sk

**Department of Avionics, Faculty of Aeronautics, Technical University of Košice, Rampová 7, 041 21 Košice, Slovak Republic, tel. +421 55 602 6140, e-mail: rudolf.andoga@tuke.sk

ABSTRACT

This paper presents the evolution of techniques and methods used for data processing, obtained from measurements in the Laboratory of Intelligent Control Systems of Aircraft Engines (LIRS LM). The task of evaluating the quality and its components (especially efficiency) is also presented. The overview shows the progress and steps that were necessary to achieve semi-automatic or automatic data collecting and processing, including respective analyses and visualizations. The data processes in LIRS LM are covered by the developed and deployed state-of-art scientific research information system (SRIS), as well as other auxiliary applications (MPM analyzer, remote monitoring, interactive wiring, etc.). With respect to the nature of the studied object (iSTC-21v turbojet engine), the system also supports knowledge discovery and allows evaluating the quality by utilizing key performance indicators (KPI), which are mostly known from economic and business sector. The currently created and deployed software equipment in LIRS LM, forms an integrated platform for both the bureaucracy (publications, form templates) and scientific work (experiments planning, measurements processing, data visualization, quality evaluation, remote monitoring).

Keywords: laboratory of intelligent control systems of aircraft engines, turbojet engine, research information system, data processing, quality evaluation, key performance indicators

1. INTRODUCTION

The laboratory of intelligent control systems of aircraft engines (LICS AE) is a place of origin of large amount of different data. The particular research object, the iSTC-21v turbojet engine, generates significant amount of data during experiments. The data may or may not be measured during its operation. The majority of measured values is of technical character and is obtained by physical sensors. Resulting measurements describe the inner state of the engine (pressures, temperatures in sections of the engine (see Fig.1), RPM, thrust, etc.), as well as the values used in feedback for control purpose (fuel flow, exhaust nozzle diameter).

The technical measured data do not describe the environment of the engine. Depending on their nature, the environmental parameters may be derived, calculated or measured during, before or after the conduction of an experiment. For the purpose of evaluating operational and identification efficiency, safety and reliability (quality in general) the technical parameters of the system are only partially sufficient for operational quality. Other components (technical, informational, time related, financial) are also involved in the quality function [1]. The size of data and the number of their distinct types require integrated collection, in order to enable analysis and provide the basis for quality evaluation. Our motivation is derived from the previous statement. Our aim was to provide an integrated platform for collecting and processing of different data obtained from the operation of iSTC-21v and from its near environment [2] [3] and enable evaluation of identification and operational quality.

The ideal way is to collect measurable data and provide additional data to fill the mosaic of the engines’ state in a specific time. An information system in general creates possibilities to combine different types of data, allows implementing algorithms for data transformation, their archiving and provides space for other data intensive operations [2] [3]. This is why we have proposed, created and deployed a scientific research information system (SRIS), the importance of which is emphasized by the complexity of observed system, as well as the evaluation (quantification) of quality.

The SRIS system in its entirety consists of several implementation tasks, each of which has to be dealt with particular care. Modules of SRIS are created in order to cover particular research processes in LICS AE [1].

2. EFFICIENCY OF COMPLEX SYSTEMS

The experimental identification appears to an observer as a complex of operations connected with logistics, measurement and control of an object. All similar systems
(compared with the iSTC-21v) are characterized by failures. If the object radiates large amounts of energy, such a system is also characteristic with environmental danger. These factors have a considerable impact on the utility (efficiency) of such a system. Every occurrence of a failure and change of its characteristic danger level, lowers the efficiency and utility value of the system (device). The extent of the efficiency loss is connected with the term “loss function”, which is given in time and has  notation [4]. For example, the loss function may represent an unacceptable difference between the real output value and a modeled output value, which results in the experiment interruption in time [5]: 
\[ \ell(y_t) = \Delta y(t) - y_m(t). \]  
(1) Let \( y_m(t) \) be the ideal output characteristic and \( \ell(t) \) the additive noise, which influences the object (its elements \( x_t \)). If \( F \) is the transformation operator of the final efficiency estimation projected into the process functions of the experimental identification, then the loss function has the following form [5]: 
\[ \ell(y_t) = F_y(x_t) + \ell(t) - F_y(x_t). \]  
(2) According to (2) the efficiency is influenced only by the emerging failures [5], which are represented by the function \( \ell(t) \). The next loss function example is the square of the difference between real and ideal values of the input characteristics of the experimental [5]: 
\[ \ell(y_t) = [F_y(x_t) + \ell(t)] - F_y(x_t)]^2. \]  
(3) Equations (1) and (2) allow the definition of the loss function of the experimental identification, using the following process [5]: 
If the given task is to estimate the efficiency of the experimental identification (\( W \)) defined by a level (\( \Lambda \)), which presents the maximal quality, then we expect the loss function to be minimal. In this case the loss function will be expressed by a mathematical probability (\( M \)) in form: 
\[ W(t,\Lambda) = \inf \left[ \ell(y_t) \right] \]  
(4) For the (4) case, the functional efficiency \( W \) will be expressed by the mathematical probability of the output error coordinate of the object in time \( t \) and for the (3) case by the mean quadratic error. 
If the condition of successful experimental identification is the quality of its level, the value of which is known beforehand, that is, the loss function does not exceed a given limit, then the criterion of functional efficiency is chosen as the probability of achieving the given limit. 
\[ W(t,\Lambda) = \inf \left[ \ell(y_t) \right] = \inf \left[ \ell(y_t) \right] \]  
(5) where \( \inf \left[ \ell(y_t) \right] \) is the achieved value of the loss function. 

The criterion of the functional efficiency has universal character. Mostly using the criterion (5) it is possible to estimate the efficiency of ergatic systems, because there is no need for physical knowledge about their functions. Moreover, it is independent of their complexity and connections with other systems. The operating systems, which are elements of the efficiency of an ergatic complex, are also a part of the efficiency of experimental identification. The mathematical probability of fulfilling the tasks of experimental identification is not always applicable, because its physical meaning and expression in form of physical units is different for various systems of a complex. 

In the previous theoretical considerations it has been stated that the evaluation of efficiency may be done using the loss function. As a substitution for the unknown parameters of the loss functions that may be difficult to measure, we may use empirically created coefficients called key performance indicators (KPI). KPI indicator is a performance indicator created according to SMART criteria (specific, measurable, attainable, realistic, time-related) [3]. Using KPIs, even a combination of measurable values may be used as a target observed parameter for evaluating the efficiency. Moreover, it is simple to plan values of a KPI and differentiate between successful and unsuccessful cases in tests. Depending on the form of the input parameters, the KPIs differ in the unit of measure. In practice, two possible cases may occur. The KPI adopts the unit of measure of the input parameter or adopts a new unit of measure. The second case happens, if the KPI represents a combination of more initial parameters with different units of measure. A special case is that the KPI cannot be assigned a unit of measure. Then the input values have to be normalized according to an interval and consider the KPI a unit less parameter. 

For cases of practical deployment, they are usually defined in form of equality or inequality in form of: 
- single parameter with interval bounds (closed interval or one side open), 
- difference of two parameters with interval bounds (closed interval or one side open), 
- ratio of two parameters with interval bounds (closed interval or one side open), 
- specific relation defined by an equation or a model, the result of which is the value of a single coefficient (or a functional value). 

The evaluation itself is done using the general figure-of-merit equation (\( W = Q/S \)) [5] that is defined as a ratio between successful KPIs and all measured KPIs in the current context. The calculated value is understood as a coefficient of success for a specific partial quality function. 

Let us assume that \( n \) KPIs of type \( ET \) have a preference value \( v_{KPI} \). Then, the specific efficiency for the \( ET \) type is calculated using the equation: 
\[ W_{ET} = \prod_{i=1}^{n} \frac{Q_{KPI}}{S_{KPI}} v_{KPI} \]  
(9) Let us consider that for all efficiency types \( ET \), where \( m \) is the number of efficiency types, a preference value \( v_{ET} \) exists. Then, the total efficiency of a single elementary research part (e.g. a test or a measurement) is calculated by: 
\[ W_A = \prod_{i=1}^{n} W_{ET} v_{KPI} \]  
(10) If the preference values in equations (9) and (10) are defined on interval \([0,1]\), then the total efficiency \( W_A \) also belongs to this interval. By transforming these equations in the similar manner (the same as (9) to (10)), we may derive the equations for solving the efficiency of higher levels of research planning (operational, tactical and strategic plans). This means that the total efficiency (quality) may be expressed as a relative success value, achieved in a specific
research step, that is, evaluate the efficiency (quality) of research. The advantage of this method is its simplicity and fast calculation of results. Its greatest advantage is the simple algorithmization, which preordains it to be implemented and deployed in the computer (program) environment.

The particular studies and possibilities of evaluating identification tests were described in student final theses [6] and [18].

Problem fields, which have been briefly described in the previous parts of this chapters depict the complexity of the tasks, as well as the object iSTC-21v itself. In publications [3], [7], [8], [9], [10], [11], [12], we described each essential elements and specified the conditions that have to be fulfilled in order to carry out the semi-automatic efficiency valuation. As one of the possible approaches to solving this complex problem, we have chosen the creation of a unifying platform (scientific research information system SRIS), which would serve to this purpose in a long-term manner and would also motivate the operating personnel to use its functions in the real-world practice. In our previously published papers, we emphasized the need for an IS not only from the efficiency valuation point of view, but also from the point of view of knowledge discovery in scientific and research work. We also stated possible implementations of suitable functions for this purpose.

3. DESCRIPTION OF THE CONCEPT OF SRIS

The structure of SRIS has been designed in such a way, so that each of its modules would cover and simplify ordinary actions done by an operator in LICS AE. It is necessary to note that the SRIS system is designed as IS with passive approach to knowledge discovery. The functions were divided into two categories according to their importance in fulfilling the given task. The tasks are defined from the current state of LICS AE.

3.1. The MPM-Analyzer application

In the first iteration, it was important to design solutions for elementary actions and test them in the real practice. The first step was the creation of a desktop testing application used for measurement analysis. Its goal was to serve as the first contact of the operator with information generated directly after conducting the measurement (later the application was given the MPM Analyzer name). This application made it possible to visualize parameters of a measurement in a two dimensional graph. The following step was the addition of other functions, which allowed to open multiple measurements and compare them, normalize measured values for “in scale” comparison of differently scaled parameters (e.g. temperature $T_{ce}$ – range 10°C to 1200°C with $P_{ce}$ – range 0 At to 2,3 At). This application has been created in C# language in .NET Framework and using the trial version of DevExpress framework. In addition to its practical usage in LICS AE, the application was used as a case study to discover requirements of operators for functionalities and the user interface. The results of this study were published in [13]. Another study connected with the usability of proposed functions in a particular environment of thermodynamic systems research was published in [14], [15]. Also based on this study, we specified functions usable and also suitable for the given research on iSTC-21v object in LICS AE laboratory.

The final version of the application (Fig. 2) contains the following functions:
- batch reading of measurements,
- basic statistics of the chosen variable,
- different types of graphs for variable plotting,
- visualization of the parameter time relation,
- visualization of two parameters,
- plotting normalized variables, their first or second derivatives or combinations (e.g. normalized first derivative)
- calculation of area under a curve using the numerical definite integration,
- clustering algorithm of measured values,
- measurement playback visualized in a dashboard form (Fig. 3).

![Fig. 2 Screenshot of MPM Analyzer (7 measurements; x-$T_{ce}$ temperature; y-RPM)](Image 289x190 to 527x320)

![Fig. 3 Example of the measurement playback screen in MPM analyzer application](Image 299x386 to 514x536)
this application was the emphasis of the fact that operational logs and experimental planning have to be implemented in the SRIS concept, so that the efficiency evaluation would be enabled.

The SRIS system should contain the following functions that resulted from the MPM Analyzer testing application:
- experiments planning,
- measurement saving and administration,
- operational log management,
- basic analytical functions,
- efficiency valuation.

Other requirements originated from basic activities of the laboratory LICSAE, which resulted in the following proposed functionalities of SRIS:
- environment settings according to LICSAE environmental conditions,
- digital library archive,
- remote monitoring of the iSTC-21v.

3.2. Design of the experiments planning methodology

In proposal of the experimental planning methodology, we used the theoretical background explained in [2] and also took the expectations for efficiency valuation mentioned in chapter 2 into account. If we consider the usage of the tree leveled hierarchical model of experimental planning for the LICSAE conditions, the time and logical allocation would be as follows:
- strategic plans – 6 months to several years – research level,
- tactical plans – 1 week to 6 months – research tasks level (stages),
- operative plans – 1 day – test series level.

Using the mentioned planning structure, it is possible to evaluate the efficiency for a particular test, test series, research task or entirely a research stage or the global research. On the elementary level, it is also necessary to plan the expected results, using rules represented by key performance indicators for the plans of the lowest level. Using this approach, we state the success criteria of the test, the success of which we may evaluated dichotomously and thus calculate the resulting functional efficiency of the test/plan using the classic figure-of-merit equation.

In case of other types of efficiency, we may use the analogy and evaluate the success or failure of the key performance indicator instead of the occurrence of a functional failure. The way of implementing this types of valuation was published in [2]. The proposed way of implementation is depicted in the schemas in Fig. 4 and Fig. 5, supports both the IS with active and passive approach to knowledge discovery.

![Fig. 4 Scheme of the proposed way of implementing the experimental planning in SRIS [8]](image-url)
3.3. Design of methodology for measurement saving

From the point of view of the logical placement of a measurement in the proposed planning hierarchy, the measurement is defined as an elementary test. Thus, it is accepted as a subset of the operational plan (test set). This is why all KPI indicators of an operative plan are also used in all measurement that belong to this plan.

On the other hand, the measurement is the central point of interest, so the need to aim the database layer mostly on saving the measurements and easy operations with data emerges. The way of representing the data is connected with the data transformation. During the overview of possible ways of representing the data, we considered several alternatives that would reflect the volume of data and expected speed of their processing, needed to be achieved in SRIS. The ideal solution was discovered after a practical test in [16]. The results have shown that to save data outside of a relational database will benefit the speed, as well as the next manipulation with saved data. Only a reference of the data file should be stored in the relational form. The proposed model of measurement saving (left) and reading (right) is depicted in the following scheme.

**Fig. 5** Detailed description of the way of implementing the experiment planning in SRIS [8]

**Fig. 6** Methodology of saving (left) and reading (right) of measurements in SRIS system
From the Fig. 6, the existence of a data model is clear. Its structure was defined and published in [17] based on the data character in given measurement files.

3.4. Design of the log structure

The proposed concept of laboratory logs, which is typical for LICS AE is derived from the requirements and has the following structure:

- **operational log** – is directly connected with the conducted and saved measurement. The entry to an operational log should be done in the last step of the measurement saving process, after the reference to the measurement file is saved in the database. As we already mentioned in the previous chapter, the data model contains descriptive statistical parameters, which briefly describe the operational values obtained from during the measurement. These values are saved as the entry of the operational log during the measurement saving.

- **Service log** – logs all technical modifications of software and hardware with their short description. The service entry may be of different character but always has the same interpretation. This means that the service log may have been done as a response to a technical failure or as an upgrade of the tested system but in both cases it is logged as a service action.

- **error log** – saves all errors and failures that manifested during the system’s operation and were defined beforehand.

By combining these three logs, it is possible to briefly describe the state of the laboratory but also depict the evolution of each one of the operational parameters based on failures, errors or service actions that were done (e.g. using a combined overview graph). The conceptual design of the LICS AE logs was composed and published in [3] and its simple structure is depicted in the following figure.

**Fig. 7 Scheme of the LICS AE logs in SRIS system [3]**

3.5. Proposal of analytical functions provided by SRIS

In chapter 3.1, we mentioned a list of methods, which were given to operators for the testing purpose. This list has been reduced and also supplemented by functions requested by operators. The result of these statements was the need for the following functions:

- visualization of a parameter in time,
- visualization of normalized parameters,
- playback of measurements in form of a dashboard,
- basic statistical characteristics of a measurement,
- calculation of a chosen percentile,
- calculation of the class intervals with their frequencies, calculation of relative and absolute frequencies for each one of the intervals (distribution character of the measured parameter).

3.6. Proposal of a methodology for efficiency valuation in SRIS

Based on the theoretical considerations, which were described in chapter 2, we have chosen the dichotomic success valuation of the key performance indicators as the basis for solving the efficiency evaluation problem [10]. The concept of evaluation is dependent on the type with which the KPIs are designed in the experimental planning phase. It holds that this procedure of evaluation is universal for every element of quality (according to its definition). The first step is the creation of a KPI with specific bounds of success for a given operative plan (see. Tab.1).

**Table 1 Example of a set of KPIs for a single operative plan [3]**

<table>
<thead>
<tr>
<th>KPI name</th>
<th>KPI type</th>
<th>KPI character</th>
<th>KPI bounds</th>
<th>Unit of measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximal t4c temperature</td>
<td>Efficiency</td>
<td>Technical</td>
<td>(n/a,950)</td>
<td>°C</td>
</tr>
<tr>
<td>Maximal fuel consumption</td>
<td>Efficiency</td>
<td>Technical</td>
<td>(n/a,6)</td>
<td>kg.min⁻¹</td>
</tr>
<tr>
<td>Average launch time</td>
<td>Efficiency</td>
<td>Time-related</td>
<td>(9,12)</td>
<td>s</td>
</tr>
<tr>
<td>Repairs cost</td>
<td>Efficiency</td>
<td>Financial</td>
<td>(n/a,250)</td>
<td>€</td>
</tr>
<tr>
<td>Time between failures</td>
<td>Reliability</td>
<td>Time-related</td>
<td>(6000,n/a)</td>
<td>s</td>
</tr>
</tbody>
</table>

The second step is to carry out the measurements and map the measured values to the observed KPI indicators. This means that the indicators that may be obtained directly from the measurements are filled in automatically. Those that may not be derived nor automatically obtained from measured data are entered manually by the operator. Next, it is necessary to apply the evaluation method for KPIs (figure-of-merit). The result for an example set of values is given in the Table 2 and will be equal to 0.6 (W = 3/5), which means the achieved quality of measurement is 60%.

**Table 2 Example of a set of KPIs and obtained values from a single measurement [3]**

<table>
<thead>
<tr>
<th>KPI name</th>
<th>KPI bounds</th>
<th>Measured values</th>
<th>Binary result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximal t4c temperature</td>
<td>(n/a,950)</td>
<td>780</td>
<td>1</td>
</tr>
<tr>
<td>Maximal fuel consumption</td>
<td>(n/a,6)</td>
<td>1,72</td>
<td>0</td>
</tr>
<tr>
<td>Average launch time</td>
<td>(9,12)</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Repairs cost</td>
<td>(n/a,250)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Time between failures</td>
<td>(6000,n/a)</td>
<td>7522</td>
<td>1</td>
</tr>
</tbody>
</table>
If we want to calculate only a particular element of the quality vector, then we only choose the respective KPI parameters. For example according to Table 1, and Table 2, achieved technical efficiency if 50% (\( W = 1/2 * 100 \)).

Using the aggregation of the partial results of measurements, it is possible to achieve quality values on the level of operative, tactical or strategic plans. The pilot practical implementation of the stated algorithms with relative valuation of success (upgraded algorithm), based on the position of the result in interval has been conducted in [18].

3.7. Methodology of LICS AE environment representation

In general, it is possible to represent any environmental conditions by a suitable database layer, where the enumeration tables represent categories of the system settings and the linking tables represent the relations between the enumeration tables. Based on this consideration, we created a wide database model, which enables its universal application to other laboratory objects with close character. The database model contains the following essential elements:

- entities for representing the environment state of the object (sensors, sensor types, units of measure, error lists, objects list),
- entities for representing the error states of the object (enumeration tables with three levels of errors with given error codes and description),
- entities describing the type and character of saved data (enumeration tables of characters and types of quality, characters and types of the linked variable, statistical characteristics)
- entities describing key performance indicators,
- entity with information about researchers.

The detailed meaning of each of the enumeration tables in mentioned in the system manual of SRIS.

3.8. Concept proposal of the remote operation monitoring of iSTC-21v

One of the aims of the future research project is to develop a wide platform for virtualization and remote administration of the laboratory aperture. Firstly the development of passive applications with one directional flow of data are created (remote monitoring). In the second phase, the active functions linked with the remote calibration of sensors, operational parameters settings and wide-range data broadcasting, will be done. We considered important to also address these tasks in this early stage because of the possible implementation of the remote functions in the SRIS structure later [19]. Whilst accepting the constraints and the dynamic growth of the current web technologies, we proposed a structure of transfer, which is platform independent and may be implemented in different environments using available application frameworks [3].

Development and deployment of this module is technically difficult and time consuming and is not directly related to the thematic content of this thesis. Because of these reasons, the module will be fully created based on the proposed concept in the future as one of the outputs of a current research project.

The current proof of concept result is presented as a simple client desktop application for data visualization in real time (low-latency) and a simple UDP sender module in LabView for one directional data transfer. The desktop application created a virtual dashboard based on an initialization string and showed the data stream after the initialization has been successfully done. A single data stream element (data string) consisted of the names of the sensing cards, sensor names and the measured value in a particular time. A screenshot of this application is shown in Fig. 8.

4. CONCLUSIONS

This paper describes a full proposal of a scientific research information system. Its modules may be adapted to be used in different situations in different research tasks. Its current implementation in LICS AE allows to take care of research agenda, process measured data, as well as allows to evaluate the efficiency of the research and its tasks. The database has been designed according to specific needs for these types of applications but mostly for the specific approaches to data reading and writing. The KPI indicators have been implemented in a nontraditional environment as a solution for evaluating efficiency. We proved that KPIs may also be well adopted in research information systems and their environment, not only traditionally in technical information systems and managerial information systems.

ACKNOWLEDGMENTS

The creation of this paper was supported by KEGA No. 014TUKE-4/2015: Digitalization, virtualization and testing of a small turbojet engine and its elements using stands for the purpose of modern applied lecturing.

REFERENCES


BIographies

Vladimír Gašpar graduated (MEng) with distinction at the department of Cybernetics and Artificial Intelligence. He is currently an assistant professor at Department of Cybernetics and Artificial Intelligence. His current research is aimed at mobile and embedded applications for stream processing and mining. The topic of his dissertation was to solve the problem of evaluating quality of experimental identification and operational quality of complex thermodynamic systems. He is currently involved in domestic research projects in fields of knowledge discovery, data mining, complex systems analysis and efficiency analysis.

Rudolf Andoga graduated (MEng) with distinction at the department of Cybernetics and Artificial Intelligence. He is currently an associate professor at department of Avionics, Faculty of Aeronautics at Technical University of Košice. The topics of his research include intelligent control of aircraft engines, automatic flight control systems, reliability and efficiency of full authority digital engine control systems (FADEC), modelling and diagnostics of complex systems.