Uncertainty analysis in the prediction of human operator violation using neural networks

Z. ZHANG, F. VANDERHAEGEN, P. MILLOT
LAMIH - UMR 8530 CNRS
Laboratoire d'Automatique, de Mécanique et d'Informatique industrielles et Humaines
University of Valenciennes, Le Mont Houy
59313 Valenciennes Cedex 9 - FRANCE

Abstract: - This paper contributes to the analysis and the prediction by the Artificial Neural Networks (ANN), taking into account uncertainty, of the deviated intentional behaviours of the human operators in the Human-machine systems. This type of behaviours is a particular violation called Barrier Removal. The objective of our work is to propose a predictive approach of the Barrier Removal by considering a multi-reference, multi-factor and multi-criteria based evaluation. Human operator's evaluation can be uncertain. Uncertainty on their subjective judgements is integrated in the prediction of the Barrier Removal. We then validate the proposed approach by a railway application within the framework of a European project Urban Guided Transport Management System (UGTMS).

Key-Words: - Artificial neural networks, prediction, uncertainty, human-machine system, human reliability, violation, barrier removal.

1. Introduction

As it's well known, more than 70% of accidents or incidents in complex industrial systems are due to incorrect human behaviours [1]. The main present challenge when designing future human-machine system is to focus on the human behaviour study to define more reliable system taking into account two important compromises:

– Despite the possibility to design an entire automated system, human operators may be maintained into the control and supervisory loop of the process in order to avoid a loss of human expertise when the automated system fails.
– Despite a permanent presence of human operators into the control and supervisory loop of the process, they may be fallible in case of the occurrences of particular control situations, such as overloaded, urgent or degraded situations.

Human factor study may take an important place when designing future complex Human-machine System (HMS). Although human errors have received close attention from psychologists and others for well over a century, the study of intentional violations is still in its early stages [2]. The importance of industrial safety violations increased after the Chernobyl accident. According to J. Reason, five of the seven human actions that led directly to the accident were deliberated deviations from written rules and instructions, rather than slips, lapses or mistakes. As a matter of fact, the violations have been mentioned in a number of contexts, however research on violations is still insignificant in comparison with what is known about slips, lapses, and rule & knowledge-based mistakes [2].

This paper focuses on an approach on the analysis and the prediction by the artificial neural networks, taking into account uncertainty, of the deviated intentional behaviours (violation) of the human operators in the HMI. The first part presents three evaluation aspects of the human operator actions in complex HMS. The second part proposes uncertainty-based prediction approach of BR using neural networks. This approach is validated by a railway simulator experiment using the BR experimentation data, and the third part gives the preliminary experimental results. In the final section, we draw our conclusions from this study and offer perspectives for future research.

2. Evaluation aspects of the human operator actions in complex HMS

An action of a human operator is initially delimited by several borderlines such as the limit of cost acceptability, the limit of available resources and the limit of the prescribed safety acceptability (Fig. 1) [3,4,5].

The level of possible migration depends on the maximum functional limit under which the users of a given machine might accept the safety. Amalberti tries an explanation to the production of some violations through the need for the human operators to manage a compromise between three joint objectives sometimes is in contradiction: performance imposed by the organization or by the human himself, safety of the system and of the operator, and the cognitive and physiological costs of theses activities (workload, stress etc.) [6]. These 3 dimensions are limited and they bound the human action field [3]. An action which crosses the limit can result in a loss of control and in an incident or an accident. This type of action is named as Barrier Removal (BR) in the paper. It concerns the particular violations that are made without any intention to subjectively damage the HMS. Divergences of risk acceptability may then appear between the different references, e.g. the designers of a machine and their users.

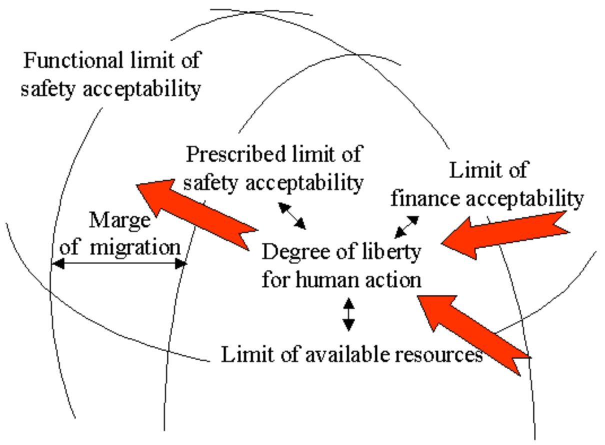

Fig. 1 Degrees of operational migration [4]

Multi-reference based BR evaluation

There are often differences between the task prescribed by the designer and the actual activity in its operational context, due to a variety of individual, technical and/or environmental factors. For the designers of a HMS, the objective of risk analysis is usually limited to an assessment of risk on safety. It is then a mono-criterion process. Its validation stops evolving when the machine is on field operation and is quite stable because it concerns a common decision. Nevertheless, this process of risk evaluation is external because it is made independently of the users and is limited to the technical failures.

Concerning the users, they have to control risks associated to operational situations by evaluating them after their detection and by intervening on the piloted process to avoid the occurrence or to limit the consequences of a given event. Each operational BR is determined by different motivational factors.

Multi-factor based BR evaluation

When a barrier is to be removed, both the positive and the negative consequences of its removal should be taken into account:

– The immediate cost of removal: In order to remove a barrier, the human operator must sometimes modify the material structure and/or the operational mode, which usually leads to an increased workload and can have negative consequences on productivity or quality.
– The potential deficit: Because removing a barrier introduces a potentially dangerous situation, such action creates a potential deficit, due to the related risk.
– The expected benefit: Barrier Removal is a goal-driven behavior seen to offer an immediate benefit that outweighs the cost.

Multi-criterion based BR evaluation

The operational BR is much more a multi-criterion risk control process. This control is multi-criterion because it takes into account not only the system safety but also economical criteria such as production, quality or social criteria such as motivation or workload. Depending on the variability of the operational situations to be controlled and on the inter-individual and intra-individual differences, the risk control process is dynamic and variable. Moreover, it can concern technical failures, human and organisational errors, violations and additional uses.

During BR analysis, the evaluations of all three factors (benefit, cost, potential deficit) are provided for each barrier class in terms of several performance criteria, which makes it complicated to directly identify the removal status of a barrier, and/or to easily group the similar BRs synthetically. Clearly, it is uneasy to capture the complex nonlinear relationships that exist between the different criteria, nor to know the similarity/proximity between all BRs.

There are two problems: first, the classification of all BRs in terms of the different performance criteria, and the identification, if possible, of the pertinent BR criteria for a given HMS by looking for or memorizing the similarity/proximity between all BRs; and second, the removal prediction for new/changed barriers, based on the identified criteria and the memorized similarity/proximity.

The Artificial Neural Networks have potential for dealing with the above-mentioned problems. A series of approaches of prediction of the Barrier

Removal using ANN have been developed to anticipate or predict with the retained criteria a removal of given barrier on the given system by considering, on the one hand a network by criterion of performance (mono-performance), and on the other hand, a network taking into account several criteria (multi-performance) [7]. It should be noted that these approaches are able to deal with not only the subjective data but also the objective ones if the latter are available. Based on these connectionist models and methods developed in our laboratory, uncertainty analysis and treatment for the subjective evaluation of human operators are further implemented and integrated in the overall methodology.

3. Uncertainty-based prediction of BR using neural networks

As stated above, BR is a safety-related violation. Its effects can be analyzed in terms of benefit, cost, and potential deficit. In order to allow designers to integrate BR into the risk analysis during the design phase or during re-design work, we have proposed three Self-Organizing Map (SOM) predictive algorithms [8]. The task that we seek to model here is the activation or the removal of barriers by the human operators.

The safety structure of any given HMS can be seen in terms of several barriers. In compliance with the pertinent regulations, standards and technical guidelines, the designers design their systems, paying particular attentions to safety concerns. They equip their systems with barriers in order to reduce human errors, limit failure propagation and/or protect human operators from technical failures.

However, in case of operational contexts, to deal with the different contexts and obtain optimal results, a series of connectionist models and methods of BR using ANN have been defined. As an artificial neural network, the Self-Organizing Map is designed originally for multidimensional data reduction with topology-preserving properties [9,10]. The proposed connectionist methods have then been validated by an experimental manipulation to analyse and/or predict with the retained criteria a removal of given barrier by considering, on the one hand a network by criterion of performance (mono-performance mode), and on the other hand, a network taking into account several criteria (multi-performance mode).

During the data collection on BR, each evaluation on each factor of BR has its uncertainty level. Factor evaluation with different uncertainty levels may have different number of the subsets. The weights should then be allocated respectively to each element of the subsets. Different weight allocation can be defined, e.g.:

– The lower the uncertainty level is, the more representative the given value is and thus the associated weight will be high.
– The lower the uncertainty level is, the less numerous the values which constitute the corresponding subset are.
– The more a value is close to the one evaluated by the human operator, the less its weight is different from the evaluated one.
– The sum of the attributed weights is equal to 1.

Combination among the subsets of all factor evaluations can then be implemented. A final format of data on BR with the consideration of uncertainty is illustrated in Fig. 2.

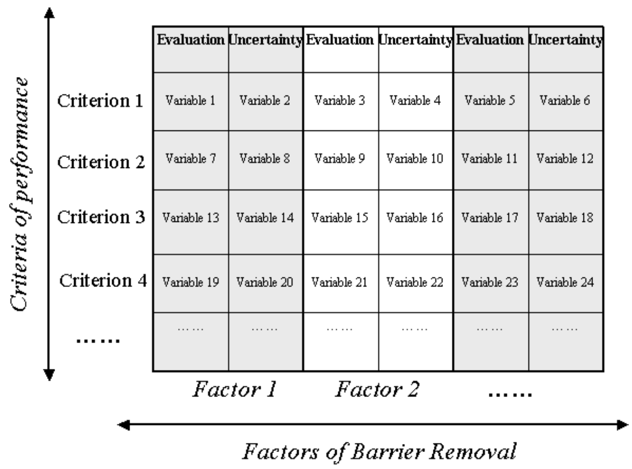

Fig. 2 A format of BR data with uncertainty evaluation

As above-mentioned, the evaluation judged by human operators on the BR factors has its uncertainty, a pre-treatment of uncertainty data is necessary. Fig. 3 shows the principle process of uncertainty-based BR prediction.

Each predictive network is composed by two parts. First part concerns the network training – BR classification:

1) In *Unsupervised Self-Organizing Map* (USOM) training, the input data are the subjective evaluations of benefit, cost and potential deficit in terms of the different performance criteria;
2) In *Supervised Self-Organizing Map* (SSOM) training, the input data are the same as those in Unsupervised SOM, but include a removal label for the corresponding barrier;

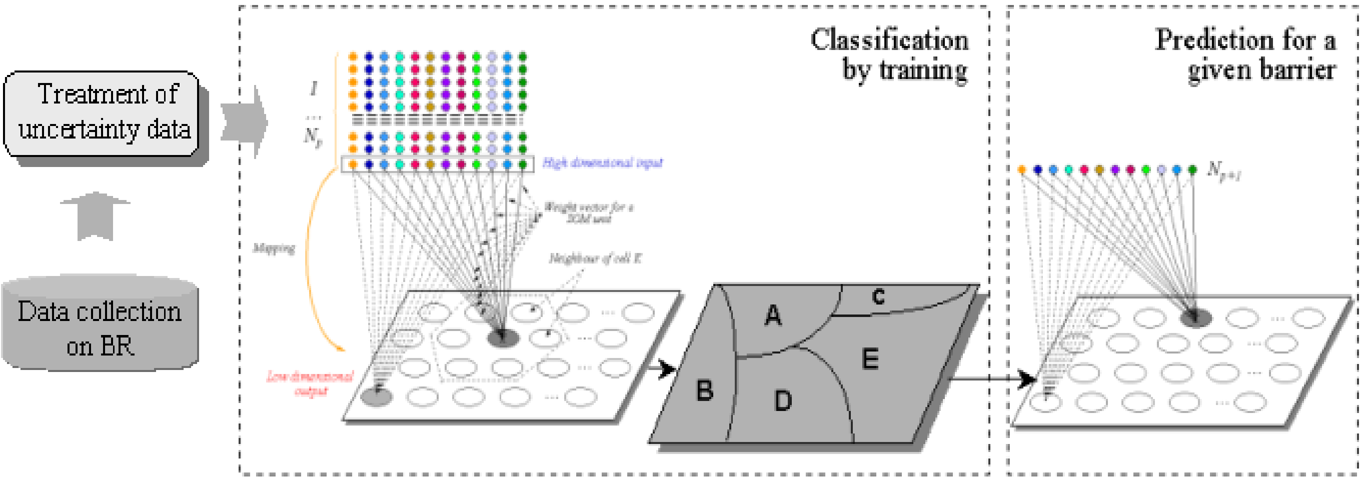

Fig. 3 Principle process of uncertainty-based BR prediction

3) In *Hierarchical Self-Organizing Map* (HSOM) training, the input data are the same as those in Supervised SOM. The network can be formed by classifying this data into parallel subsets, according to the personalities of human operators. For example, experimental BR data may be grouped into several subsets related to controller culture background (e.g. nationality).

Second part of each network is the prediction based on the identified criteria and the similarity/proximity memorized through the training process. When the target value is known, the SSOM algorithms can be used to make the classification & the prediction. When the fact that different people have different characteristics is important - e.g., when it would be helpful to group experimental BR data concerning human operators into subsets related to their personalities - the HSOM algorithms may be used. When the target value is unknown, the USOM algorithms can be used, as in data mining for example.

4. Application in railway experiment

An application of the proposed method is validated by a railway simulator experiment.

4.1 Experimental protocol

This feasibility study is based on an experimental platform called TRANSPAL (French acronym for Train Transformation System). TRANSPAL simulates train movements from a depot to another depot crossing several exchange stations on which human operators convert products placed on the stopped train [11].

A human operator has to control the train traffic flow. Several risks have been identified for such a controlled process:

– Derailment of trains: when a train is authorised to move, it may derail if the corresponding switching device is not operated correctly.
– Shunting error: a train may be directed toward an incorrect route.
– Collision between trains such as a face-to-face or an overtaking collision.
– Accident at transformation stations: an accident may occur when human operators are not aware of the movement of a train before it enters and leaves a transformation area.
– Important delay on the planning: products of trains may be treated not in time or partially.

In order to limit risks of miss-control, several barriers are proposed in order to control the traffic flow, the routes of the trains, to prevent collisions or derailments, and to inform operators at transformation areas. There are immaterial barriers such as procedures that constraint the human controllers' behaviours:

– To respect the direction of movement: a train cannot reverse except in case of an error of control or of a particular manoeuvre using the middle platform of each station.
– To put systematically the signals on red: a green signal authorises a train to move.
– To announce the train arrival and departure from transformation areas.
– To respect the timing knowing that is better to be in advance. Train movements are planned regarding the time of departure from a depot, the times of arrival into intermediate transformation stations, and the time of arrival into the destination depot.

There are materiel and functional barriers such as signals with which human controllers have to interact:

– Signals to prevent traffic problems related to the inputs and the outputs of trains from depots.
– Signals to prevent traffic problems into transformation areas: there are signals to control inputs and outputs at the transformation areas and delays to inform the treatment of the content of a train is in course.
– Signals to prevent traffic problems at the shunting device.

Human controllers can only act on the position of the switch points, the state of the signals, the announcement of traffic flow at transformation areas. Fig. 4 presents the principle experimental steps.

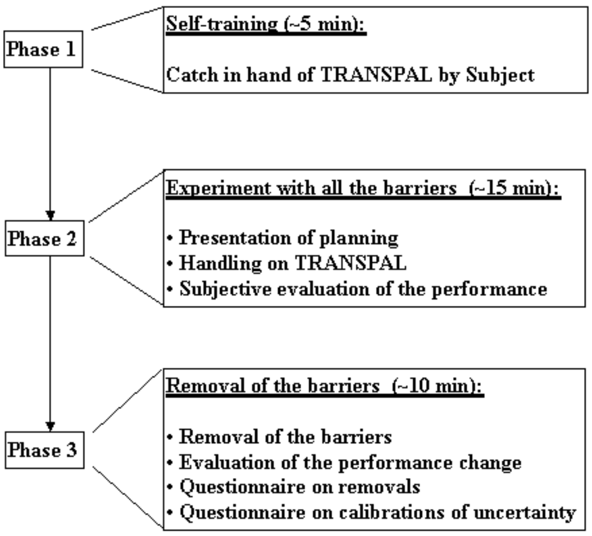

Fig. 4 Principle experimental phases

4.2 Results

Twenty human experts from a European railway projet Urban Guided Transport Management System (UGTMS) have participated in three experiment phases: the first one is a training phase to understand the TRANSPAL process and interface, the second one integrates all the operational barriers that are signals at depot, at switching device and at transformation areas, and the third one proposes to the human operators the possibility to remove some of these operational barriers. After the experiments, they answer to a questionnaire on the evaluation of the interest of the removal of class of barriers in terms of benefit, cost, and potential deficit and evaluating the associated subjective assessment certainty levels. They have to take into account four performance criteria:
– The quality related to the advancement of the planning.
– The production related to the percentage of product treated at the stations.
– The traffic safety in terms of collision, derailment and possible accident due to an incorrect synchronisation of movement announcement message at transformation stations.
– The human workload related to the occupational rate.

Results focus on the perception of the impact of the barriers of five families:
– The signals for train movements at the inputs of the depots.
– The signals for train movements at the outputs of the depots.
– The signals for train movements before crossing a shunting device.
– The signals for train movements before entering and after leaving the transformation stations.
– The signals for stopping trains at transformation areas

Before predicting human action regarding human perception data on barrier removal, a training step is required in order to determine these data distribution. The prediction phase consists assessing the action that may be performed by the human operator regarding an input vector containing uncertainty-based evaluation data on barrier removal. The input vectors of the neural network was composed by the barrier removal factors, i.e. the benefit, the cost and the potential deficit associated to the barrier removal for each performance criterion, the decision of respecting or removing the corresponding barrier, as well as the associated uncertainty level.

In order to verify the effectiveness of prediction with the uncertainty on the barrier removal factor evaluation, several cases were defined: the Case 1 corresponds to the learning phase integrating the inputs vectors of 5 human experts and the prediction is assessed for the 15 other experts. The Case 12 contains the input vectors of 16 human experts and the prediction phase concerns the last 4 human experts. The prediction rate is a comparison between the prediction given by the neural network and the real behaviour of the human experts. Fig. 5 gives an example of the impact of the quantity of learning phase input vectors on the prediction rate.

Results show that the number of input vectors used for the training phase has an impact on the convergence of the prediction rate. The prediction rate converges toward 95% using the uncertainty evaluation.

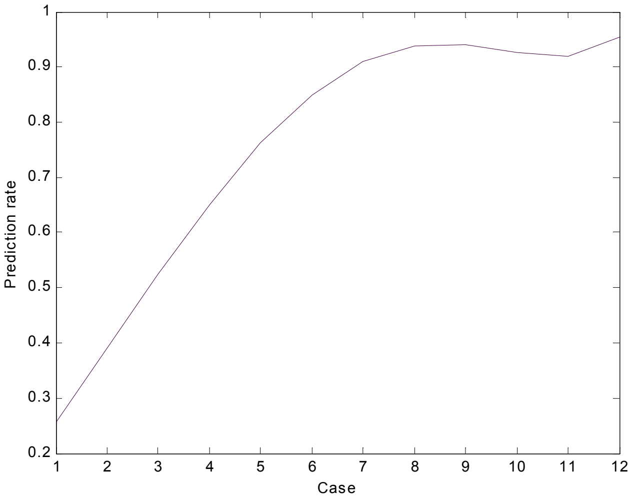

Fig. 5 Fitting curve of prediction rate with and uncertainty data

5. Conclusions

This paper has presented a method for using Artificial Neural Networks to analyze and predict BR based on the uncertainty evaluation of subjective data. Uncertainty analysis and treatment for the subjective evaluation of human operators are further implemented and integrated in the overall methodology. Representing BR factor data and its corresponding removal results as a constraint network can provide designers/users with tools that will allow them to predict with some accuracy the removal possibilities for new/changed barriers.

Retrospective analysis can be undertaken to identify or to regroup the BRs of a given HMS system using Unsupervised, Supervised or Hierarchical SOMs in such a way that the barriers most often removed can be taken into account. With sufficient training and enough sample data, the competitive neural networks can be developed/configured to predict the removal of a redesigned barrier. In the mean time, they can be used as a statistical data mining method to perform the multidimensional BR analysis.

Based on the SOM maps obtained from a training set, predictions can be made for the changed/new barriers. Thus, prospective analysis can be implemented to make the removal prediction for a new barrier. This method can also be used during the (re)design process to aid in the evaluation of existing barriers.

The proposed connectionist models and methods have been validated before by a set of experimental data without the uncertainty evaluation. They makes it possible to design a human-machine system that can identify correct behaviors from erroneous or dangerous ones, and then tolerate or use specific barriers to control/prohibit them. In the paper, the experiment with 20 railway experts in the framework of a European project shows that the proposed approach works well for the uncertainty-based BR prediction. However, the results presented here constitute only a preliminary uncertainty analysis on BR prediction. Further research is currently being done or will be done in near future. For instance, the uncertainty rate for the subjective evaluations of human operators may be treated using another allocation methods from fuzzy logic communities; the proposed approach can be further verified by comparing with another methods, e.g. the second predictive part of the proposed approach can be compared with Case based reasoning.

Acknowledgements

The paper is a part of a European 5th framework project named Urban Guided Transport Management System. The authors would like to thank the 20 experts from the project consortium and PhD candidate A. CHAALI from the laboratory LAMIH for their contributions to this project.

References

[1] Amalberti, R., The paradoxes of almost totally safe transportation systems. *Safety Science*, 37, 2001. pp. 109-126.

[2] Reason J., A system approach to organizational error. *Ergonomics*, (1995) vol 38, No.8, pp. 1708-1721.

[3] Rasmussen J., Risk management in a dynamic society : a modelling problem, *Safety Sciences*, 27, 2/3, pp.183-213, 1997.

[4] Millot, P., Human machine interface versus human machine interaction for safety of automated systems, 3rd International Conference SIAS 2003 - Safety of Industrial Automated Systems, Nancy, France October 13-15, 2003.

[5] Vanderhaegen F. (2003). *Analyse et contrôle de l'erreur humaine*, Hèrmes science, Paris, France.

[6] Amalberti R. 96, *La conduite des systèmes à risques*, PUF, 1996.

[7] Zhang Z., Polet P., Vanderhaegen F., Millot P. (2003). Artificial Neural Network for Violation Analysis. *Reliability Engineering and System Safety,* in press, July 2003.

[8] Zhang Z., Vanderhaegen F., A method integrating Self-Organizing Maps to predict the probability of Barrier Removal, Chapitre 30 in H. Bozdogan (ed.), "*Statistical Data Mining and*

Knowledge Discovery", CRC Press, New York, July 2003. pp.499-513. Selected and extended paper of the C. Warren Neel Conference on the New Frontiers of Statistical Data Mining and Knowledge Discovery, June 22-25, 2002, Knoxville, Tennessee, USA.

[9] Kohonen T., *Self-Organizing Maps*. Springer-Verlag, Third edition, Berlin, Heidelberg, Germany, 2001.

[10] Dreyfus, G. et al., *Réseaux de neurones, méthodologie et application*, Eyrolles, Paris, France. 2002.

[11] Zhang Z., Polet P., Vanderhaegen F., (2003a). Towards a qualitative predictive model of violation in transportation industry, European Safety and Reliability Conference Esrel2003, Maastricht, The Netherlands, June 15-18, 2003. pp.1743-1752.