Fuzzy Harmonic Systems for Traffic Risk

Walter Bel[1], Daniela López de Luise[1],

[1] Universidad Autónoma de Entre Ríos (UADER),
Concepción del Uruguay, Entre Ríos, Argentina
bel.walter@uader.edu.ar, daniela_ldl@ieee.com

Abstract. This paper aims to present a model for traffic risk prediction. Its contribution is the adaptation of Fuzzy Logic applied to Harmonic Systems in order to make it more flexible and powerful in certain contexts. The possibility of having a good traffic risk prediction opens a practical possibility to successfully improve the security not only for drivers, but also for pedestrians and cyclists. The proposed model is able to process in real-time with simple data provided by the environment and the individual whose risk is being processed. The scope of this paper covers the technical description of the model, statistical analysis and comparison with alternates using Traditional Ruled-Expert Systems (RES), Harmonic Systems (HS) and Fuzzy Harmonic Systems (FHS). Also a short proposal for the prototype is described. Results indicate a remarkable improvement for the FHS predictor compared to RES and HS.

Keywords: Fuzzy Harmonic Systems, Harmonic Systems, Expert Systems, Traffic Risk Prediction

1 Introduction

With the development of the digital technology, new data sources are capable of generating great diversity and quantity of data. In accordance to the International Data Corporation (IDC) today data are increased by 50% per year [1]. These characteristics, when it comes to important numbers of records, they become unmanageable for traditional systems. It is thus that the techniques of big data and deep learning take place in these circumstances, generating the possibility of enhancing less procedural algorithms in pursuit of modern heuristics. All this is at the cost of a process of increasing the bias of the model, but with the convenience of becoming manageable problematic situations so far difficult until now difficult to solve.

The number of drivers on the roads and streets has increased dramatically over the years. In accordance to the Wards Auto [2] the number of vehicles in circulation in the world went from 980 million in 2009 to 1015 million vehicles in 2010, representing a 3,57% in just one year. The traffic generated as a consequence causes congestions and complications that increase over time, in the detriment of not only of the drivers themselves but also of the rest of the cohabiting community. One of the most compelling problems in the world are traffic accidents and their consequences [3], [4] involving material losses, people injured and even multiple deaths [5], [6]. An effective risk assessment should include information of the environment, proper context information, vulnerability of the individual, sudden biomechanical forces [7], among other variables. There are many studies and statistics with models that partially represent the apparent risk, but they are not accurate predictors because they miss a representative number of hidden factors, which are indirect or partial cause. Examples of such uncovered tips are: the fact that pedestrian injuries increase with vehicle speed, sidewalk status, availability of suitable crossing facilities, lack of proper pedestrian crossings, number of lanes to cross, complexity of intersections, etc.

Among the items that are directly involved there are some that are typically considered, such as the use of seatbelt and airbag in vehicles. However there are other extra facts that relate to the impossibility of preventing risk situations (for example: sudden movements from other vehicles or pedestrians, animals crossing the sidewalk, etc.). The reaction time in such cases is often insufficient for appropriate risk nullification. On average, a person reacts in 1.5 seconds to a danger stimulus. This should be added to other facts that can slow down or alter this time making things worse. Some proposals try to generate some extra time, so the driver can limit the risk of accident, one example is the VANETS navigation platform [8].

There are proposals that are based on the use of certain technologies. By case, designing networks that can act in different environments (streets and routes) [9]. There are also works on communication between vehicles (cell phone, Bluetooth, RF, infrared, etc.) [10], [11], [12], [13], [14] as an element of additional support and prevention. In these proposals the vehicle communicates with two others (front and behind) while driving in its lane. But their success is limited and they are very sensitive to the weather conditions. Other proposals present the possibility of being coupled to the software of a GPS navigator or cell phone devices of personal use.

From Data Mining (DM) perspective, risk can be thought as a derivation of a set of variables heuristically selected as the most likely cause of some described accident. The HS are a type of mining centered on the rhythm, accelerations, static periods and other aspects related to the time characteristics of the selected patterns. HS also allows real-time processing. It is prepared for applications that require rapid reactions to data collected with certain criteria (for example a driver or pedestrian who compiles environmental data during its movement from one point to another) [15].

The FHS model extends HS with fuzzy predictions, enabling the semantic conception of the variables, thus allowing a better focusing of the predictor behavior. Instead of recognizing patterns configuration of certain parameters deterministically, it analysis just timings behavior at intervals. Taking that step in inferential capacity, it increases the bias of this model. This model changes the pattern of resonance parameters as processed by HS incorporating a fuzzy configuration. This way, according to the Occam Razzor theorem [16], a higher pattern detection power is expected in new situations. The rest of this paper presents the Fuzzy Harmonic System model (section 2), a statistical analysis and comparison of ES, HS and FHS models (section 3), the proposal of prototype FHS implementation (section 4) and finally conclusions and future works (section 5).

2 Fuzzy Harmonic Systems Model

This section presents a technical description of the FHS [17].

FHS Goal. FHS systems can be applied to cases where the problem to be treated has two characteristics: response requirements in real time and the data are not discrete, also requiring a complex context of such data. Through the use of fuzzy patterns it is possible to detect implicit knowledge in the variables defining membership functions (for example low visibility, high speed, sunset and dawn hours, etc.), and the use of a word bag implementation for categorical attributes (for example bag_rain = {Rain, Drizzle, Hail, Thunderstorm, Precipitation, Overcast, Squalls}). Fuzzy patterns can be parameterized deterministically, by set of values or by intervals. This way it is possible to use FHS in problems whose data need to be interpreted with semantic sense, before being converted and evaluated as numbers.

FHS Problems. FHS can be used in the same types of problems as HS, provided that specific patterns of interest are properly defined (for example for problems like intrusion detection in a network, production failures, Decision Support Systems, process control, etc.). Additionally, it is possible to react upon changes of behavior in those patterns. It is important to note that the focus of the system emphasizes the time of occurrence of events, not the event itself.

FHS Functioning. FHS extend the HS model approach. A deeper description of HS is out of the scope of this paper but readers interested in details may find them in [18], and for more details about HS applied to traffic risk in [19], [20], [21], [22] y [23]. The main feature that is being added at the present model is the use of fuzzy patterns and certain basic expert rules as shown in fig. 1, providing the system with the ability to contextualize variables, and incorporating commonly implicit knowledge of the context in the form of membership functions. Through a set of specifically designed operators, it is possible to manipulate logical and subjective concepts of the problem through numbers. The result of the whole process is a weighted collection of non-numeric parameters, enriched with the contribution of subjective information not considered by other approaches.

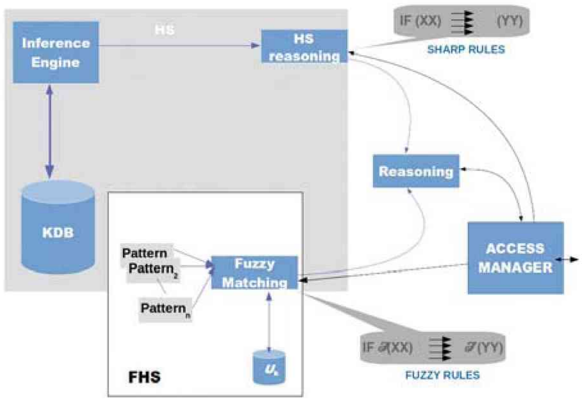

Fig. 1. FHS Model.

3 Statistical Analysis

Statistical analysis compares performance of a prototype implementing ES, HS (presented in [17], [22]) and a functional prototype that partially implements the FHS core. In order to be able to perform the comparison of the different models it is mandatory to test same cases.

-Dataset, is a collection of real world traffic at "Concepción del Uruguay" city (Entre Ríos province), these data are in the public domain (can be accessed from the link: https://drive.google.com/file/d/0B97BXscx2_9mSnpsZnB3Qm0wQnM/view?usp=sharing), fig. 2 shows a map of the city and its risk zones.

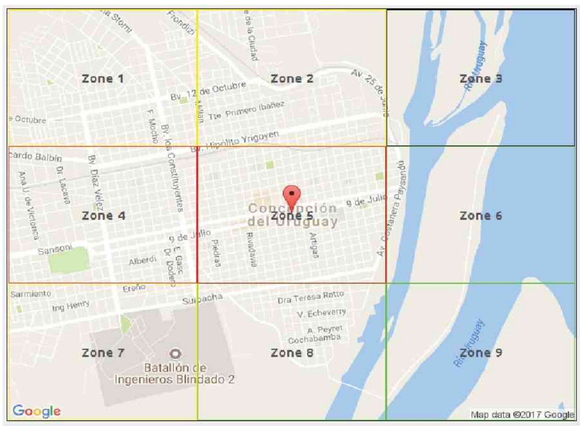

Fig. 2. Map of risk zones for Concepción del Uruguay.

As can be seen, the map has a mix of risks. Each cell represents a zone, described in the database as a tuple IDzone, two points of latitude and longitude, and risk. IDzone (identification code for the cell), its geolocation points and the level of risk associated with it (values can be HIGH, MEDIUM, or LOW). The map represents risk levels with green (low), yellow (medium), red (high).

-Patterns and rules, table 1 lists the 14 patterns used for testing HS and FHS systems. ES rules are the same as the ones described for HS patterns.

Table 1. Description of driver and pedestrian patterns for HS and FHS models.

ID	Patterns HS	Patterns FHS
Driver Patterns		
1	Type = Bike, Alcohol >= 0.2 Color = black, Belt/Helmet = false Time > Sunset	type = Bike, alcohol1 = high color = risk, belt_helmet = false time = nightfall
2	Type = Motorcycle, Alcohol >= 0.2 Weather = Rain	type = Motorcycle, alcohol1 = high weather = rain
3	Type = Motorcycle, Alcohol >= 0.2 Weather = Rain, Color = black	type = Motorcycle, alcohol1 = high weather = rain, color = risk
4	Type = Motorcycle, Belt/Helmet = false, Speed > 40, Time > Sunset	type = Motorcycle, belt_helmet = false speed1 = medium, time = nightfall
5	Alcohol >= 0.45, Color = black Time > Sunset	alcohol = high, color = risk time = nightfall
6	Gender = Male, 22 >= Age >= 18 Day = Saturday, 12:00:00 > Time > 00:00:00	gender = Male, age = young day = Saturday time1 = dawn-md
7	Gender = Male, 22 >= Age >= 18 Day = Friday 23:59:59 > Time > 18:00:00	gender = Male, age = young day = Friday time = nightfall
8	Age >= 55, Time > 18:00:00 Time < 23:59:59	age = old, time = nightfall
9	Visibility <= 4, Color = black Time > Sunset	visibility = little, color = risk time = nightfall
10	Visibility <= 4, Color = black Time < Sunrise	visibility = little, color = risk time = dawn
11	Type = Big Truck, Visibility <= 4 Speed > 80	type = big, visibility = little speed = high
Pedestrian Patterns		
12	Weather = Rain, HeadPhone = true Time > 18:00:00	weather = rain, headphone = true time = nightfall
13	Alcohol > 0.45, Time > Sunset	alcohol = high, time = nightfall
14	Alcohol > 0.3, Time < Sunrise Age >= 50, Visibility <= 4 Temperature <= 10	alcohol1 = high, time = dawn age1 = adult-old, visibility = little temperature = low

For more details about fuzzy sets used for FHS patterns and fuzzification process see [17]. Table 2 shows the initial configuration values used for HS and FHS models.

Table 2. Initial configuration of HS and FSH patterns.

Parameters	Value
u - threshold	0.03
n+ - number of system resonance cases	0
nt - total cases treated by the pattern	0
nu - user threshold inertia coefficient	0.3
n - learning coefficient	0.05
nc - distribution threshold	80
li - average time of stamp i of the pattern	0
ti - time of last activation of the pattern	0

-Results, Table 3 presents a summary of the results of the testing process for the three prediction models. The number of cases studied, instances of risk, predicted risk level (risk values can be 0=NO RISK, 1=LOW RISK, 2=MEDIUM RISK, 3 = HIGH RISK), and the cases where the zone's risk level influences the risk level calculated by the model [23].

Table 3. Summary of systems testing.

Number of cases	Description
516488	Total testing cases
15440	Cases of risks detected
3036	Cases in which the risk of the zone influence in the risk defined

Table 4 presents a detailed description of the results obtained from testing, grouped according to the main characteristics of the behavior presented by ES, HS and FHS models. It can be observed similarities and differences of the risk obtained, the number of cases corresponding to each group and subgroup, and some remarks.

Table 4. Detailed description of test results.

ES	HS	FHS	Amount	Sub-Amount	Remarks
0	0	0	501048	501048	None of the three models detect a risk situation in the case.
1	1	1	309	2018	The three models return the same level of risk to the case.
2	2	2	890		
3	3	3	819		
0	0	1	1217	7847	FHS detects risk situations of different levels. HS and ES do not detect risk for the case. This difference is due to the flexibility of fuzzy patterns.
0	0	2	4035		
0	0	3	2595		
0	1	1	4	699	HS and FHS detect the same level of risk and ES does not detect a risk situation for the case. This is because both models focus on time and not on the event.
0	2	2	398		
0	3	3	297		

0	1	0	21	35	The HS detects risk situations of different levels. FHS and ES do not detect risk for the case. FHS does not detect them due to the demanding threshold of acceptance (only resonates in cases of marked risk).
0	2	0	11		
0	3	0	3		
1	2	2	703	1868	The difference in level is due to the fact that both HS and FHS add to the level of risk the risk inference of the zone.
2	3	3	1165		
1	0	0	65	710	The ES detects a risk situation of different levels. HS and FHS do not detect a risk situation. This is due to the acceptance threshold.
2	0	0	177		
3	0	0	468		
1	0	2	17	60	FHS and ES detect a risk situation but HS does not (this is due to the acceptance threshold), the difference in level is due to the inference of the risk of the zone.
2	0	3	43		
1	2	0	324	1108	HS and ES detect a risk situation but FHS does not (this is due to the acceptance threshold), the level difference is due to the inference of the risk of the zone.
2	3	0	784		
1	0	1	11	70	FHS and ES detect a risk situation but HS does not (this is due to the acceptance threshold).
2	0	2	51		
3	0	3	8		
1	1	0	113	1025	HS and ES detect risk situation but FHS does not (this is due to the acceptance threshold).
2	2	0	766		
3	3	0	146		

From the results, a random and significant sample is extracted to analyze the relation between the prediction results of the models. Contingency tables 5 and 6 show an interesting difference between the different predictions.

Table 5. Contingency table es_predict * hs_predict.

		HS_predict				**Total**
		0	1	2	3	
es_predict	0	6623	0	209	171	7003
	1	48	219	560	0	827
	2	147	0	777	931	1855
	3	246	0	0	70	316
	Total	7064	219	1546	1172	10001

Table 6. Contingency table es_predict * FHS_predict.

		FHS_predict				**Total**
		0	1	2	3	
es_predict	0	2505	675	2405	1418	7003
	1	259	167	401	0	827
	2	925	0	390	540	1855
	3	316	0	0	0	316
	Total	4005	842	3196	1958	10001

4 Proposal Prototype FHS

As mentioned in previous sections, the Kronos prototype partially implements the FHS model. Since the goal here is to validate its improvement of efficiency with respect to the HS model, it is not required a full implementation. The prototype is coded Java for its simplicity and flexibility to migrate to Android with minimal modifications. The global design of the system allows to interact via the Internet (with off-line support for cases where there is no Internet service). There is also the possibility to use it with several devices and alternate information systems (see fig. 3).

Fig. 3. Architecture of the prototype Kronos Mobile.

Kronos Admin (1). Management of patterns, maps and zones of risk, in accordance to the functionality previously described [17], [23].
Firebase (2). Real-time database for persistence of system data and data generated by it, with support for authentication and user management. It can be used also off-line.
Kronos Mobile (3). Mobile prototype that implements several functionalities.

a. **user profile and vehicles**, allows the configuration and profiling of users and their vehicles.
b. **blood alcohol model**, based on the Widmark formula [24] using UBE (Standard Drinking Unit). It helps to determine alcohol level in blood, and infers its decreasing behavior after time t.
c. **weather model (5)**, based on the descriptions and needs determined in [17]. It is updated by means of a difference function after a time t, if the weather data present a significant difference from those obtained previously.

d. **FHS model**, to calculate the level of risk of the user and issue an alert in case of being exposed to a possible risk situation.
e. **speed calculation model (4)**, based on GPS geolocation, obtains the speed of the user's movement using the Haversine formula [25].

5 Conclusion and Future Work

This paper presents the FHS model, with a prototype that implements it. It is presented also a statistical comparison using ES, HS and a partial implementation of FHS. The characteristics of each model were presented: rules, patterns and basic semantic information about the zones used to infer the risk levels. The comparison between the different models shows interesting differences in the number of cases detected by FHS on HS and ES. This is due to the flexibility granted by fuzzy patterns and the ability to contextualize the variables and obtain subtle information.

As future work remains the full implementation of the FHS model and to perform real-time evaluations in different geographical locations of the country in order to measure the robustness of the model behavior in different situations.

References

1. International Data Corporation. https://www.idc.com/ (2017)
2. Wards Auto. http://wardsauto.com/ (2017)
3. D. López De Luise, MLW and bilingualism. Adv. Research and Trends in New Tech., Software, Human-Computer Interaction, and Communicability. IGI Global. USA. (2013)
4. O. Wash. Assessing pedestrian risk locations: a case study of WSDOT efforts. Department of Transportation. Washington State Library. Electronic State Publications. (1998)
5. Org. for Economic Co-operation and Development (OECD). http://www.oecd.org. (2014)
6. D. Alex Quistberg, J. Jaime Miranda, Beth Ebel. Reducing pedestrian deaths and injuries due to road traffic injuries in Peru. Rev. Peruana de Medicina Experimental y Salud Pública. Rev Peru Med Exp Salud Pública vol. 27 n. 2 Lima Apr/Jun. ISSN 1726-4634. (2010)
7. J. Oxley. Improving Pedestrian Safety. Curtin - Monash Accident Research Centre. Fact Sheet No. 6. (2004)
8. VANETs – A platform for the future Intelligent Transport System (ITS). Ms. Dahlia Sam. CSE (2014)
9. Vehicular ad hoc networks (VANETs): status, results,and challenges." Telecommunication Systems - Zeadally, Hunt, et al. (2010)
10. H. Fujii, K. Seki, and N. Nakagata, "Experimental research on protocol of inter-vehicle communication for vehicle control and driver support," 2nd World Congress on Intelligent Transport System. 1995, Yokohama, Japan, Nov. 9–11, pp. 1600–1605. (1995)
11. I. Sasaki, T. Hirayama, and T. Hatsuda, "Vehicle information network based on inter-vehicle communication by laser beam injection and retro reflection techniques," IEEE Vehicle Navigation and Information System Conference. Japan , pp. 165–169. (1994)

12. K. Mizui, M. Uchida, and M. Nakagawa, "Vehicle to vehicle communication and ranging system using spread spectrum technique" IEEE Vehicle Navigation and Information Syst. Conf. '94, Yokohama, Japan, Aug. 31–Sept. 2, pp. 153– 158. (1994)
13. D.W.Kremer, D. Hubner, S. Holf, T. Benz, and W. Schafer,"Computer aided design and evaluation of mobile radio local area networks in RTI/IVHS environments," IEEE J. Select. Areas Communication, vol 11, April. (1993)
14. J.-M. Valade, "Vehicle to vehicle communications: Experimental results and implementation perspectives,"2nd World Congress on Intelligent Transport Syst. '95, Yokohama, Japan, Nov. 9–11, pp. 1606–1613. (1995)
15. Daniela López De Luise, Walter Bel - "Cálculo de Riesgo en Tráfico y Peatón usando Sistemas Armónicos" - 2017, 978-3-639-53739-0. Editorial Académica Española. (2017)
16. Data Mining and Knowledge Discovery, Domingos, P. Data Mining and Knowledge Discovery 3: 409. doi:10.1023/A:1009868929893. (1999)
17. Parametric Prediction Model using Expert System and Fuzzy Harmonic System. W. Bel 1 , D. López De Luise 2 , A. Asmis 3 , D. Mansilla 4. SOFA. (2016)
18. D. López De Luise. Harmonics Systems for Time Mining. International Journal of Modern Engineering Research (IJMER). Vol. 3, Issue. 6. 3 pp-2719-2727. ISSN: 2249-6645. (2013)
19. A. Celayeta, C. Paredes, D. López De Luise, W. Bel. Traffic and Pedestrian risk evaluation with Harmonic Systems.ARGENCON. (2014)
20. D. Mansilla, D. López De Luise, W. Bel. Un Modelo con Conocimiento Experto y Sistemas Armónicos para Evaluación de Riesgos de Tráfico. EnIDI Argentina. (2015)
21. I. Acuña, E. García, D. López De Luise, C. Paredes, A. Celayeta, M.Sandillú, W. Bel. Traffic & Pedestrian risk inference using Harmonic Systems. SOFA. Romania. (2014)
22. Risk Prediction on Time with GPS Information . Daniela López de Luise, Walter Bel, Diego Mansilla, Rigoberto Malca la Rosa. "Information Technology and Computational Physics", P. Kulczycki, L.T. Kóczy, R. Mesiar, J. Kacprzyk (Eds.), Springer (2017)
23. Predicción de Riesgo basado en tiempo y patrones GPS. Daniela López de Luise, Walter Bel, Diego Mansilla, Rigoberto Malca la Rosa. ARGENCON. (2016)
24. Widmark - "Principles and Applications of Medicolegal Alcohol Determination" - Biomedical Publications - 978-0931890079. (1981)
25. Prof. Nitin R.Chopde, Mr. Mangesh K. Nichat - "Landmark Based Shortest Path Detection by Using A* and Haversine Formula" - IJIRCCE Vol. 1, Issue 2, April - 2320–980. (2013)