Risk Prediction Based on Time and GPS Patterns

Daniela López De Luise, Walter Bel, Diego Mansilla, Alberto Lobatos, Lucía Blanc and Rigoberto Malca la Rosa

Abstract Traffic produces not only pollution but also many incidents resulting in material lost and human injuries or even persons dead. But not all the incidents involve two cars, it may be pedestrian and any type of cycles (motorcycles, bicycles, tricycles, etc.). Most of the approaches try to model traffic accidents using traditional information and avoiding others, such as environmental elements, driver profile, weather, regulations, eventual circumstances like strikes with roadblocks, street reparations, railroads crossings, etc. This paper presents a model for risk prediction, and the impact of varying geographical information details on the precision of the underlying Inference System (a Soft Computing model with a ruled Expert System and a Harmonic System focused on time patterns of events). Its flexibility and robustness has a price: certainly minimal to apriori knowledge. This work outlines the working model implemented as a prototype named KRONOS, and a statistical evaluation of its sensibility to dynamic GPS information. Traffic risk requires this type of flexible and adaptive model due to the high number of alternatives to consider. The model would also be improved by adding certain specific Fuzzy Logic for pattern management during the matching process. The model would also be improved by adding certain specific Fuzzy Logic for pattern management during the matching process.

Keywords Risk prediction • Time mining • Machine learning • Expert systems • Harmonic systems • Pedestrian risk • Traffic risk

---

D. López De Luise (✉) · R. Malca la Rosa
Computational Intelligence and Information Systems Lab, Ciudad de Buenos Aires, Argentina
e-mail: daniela_ldl@ieee.org

R. Malca la Rosa
e-mail: daniela_ldl@ieee.org

W. Bel · D. Mansilla · A. Lobatos · L. Blanc
Universidad Autónoma de Entre Ríos, Concepción Del Uruguay, Argentina

P. Kulczycki et al. (eds.), *Information Technology and Computational Physics*, Advances in Intelligent Systems and Computing 462,
DOI 10.1007/978-3-319-44260-0_7

1 Introduction

One of the most compelling problems in the current world is traffic accidents, and the consequences [1, 2] involve materials lost, injuries and even death [3, 4]. But an effective evaluation of the risk must include information on the environment, context details, vulnerability of the individual, biomechanical resistance to sudden forces [5], etc.

Although there are many studies and statistics, any model representing this type of risk become apparent unless it covers a representative number of hidden factors that are indirect cause or bias. For instance, pedestrian injuries increase with a higher speed of traffic, status of footpath, availability of adequate crossing facilities, pedestrian crossing opportunities, number of lanes to cross, complexity of traffic movements at intersections, etc.

Furthermore, current and past statistics and proposals are statically defined in advance, considering just most meaningful and logical variables, but there is a lack of flexibility to append new factors dynamically. This is important, because technology evolves, society changes and therefore variables change as well. Just to mention a few of the variables, there may be the age of the pedestrians [6], subtle tips like crowd management [6], pedestrian attitudes [7], pedestrian crossings [8], etc. The prototype in this paper was tested with an initial knowledge that combine, many of these mentioned items and others. A few traffic risk models are in this line, like the Traffic Management Hazard Identification & Risk Assessment Control Form [9], that checks relevant causes and related events and G20/OECD, that is a framework for risk assessment [3]. Other proposals are still waiting for implementation and evaluation.

For pedestrian risk there are also some further alternatives. Among others can be mentioned the proposal to assess the risk of collision related to a pedestrian-based scenario [10], a Case-control approach [11], Micro-simulation Model with SSAM (Surrogate Safety Assessment Model, developed by FHWA, US) [12], a tailor-made statistical tool [13], Journey Risk Management [14], etc.

The authors in [15, 16] suggests modifying the physical environment, but to be aware of how this should be accomplished it is necessary to understand better which are all the main factors in most of the cases.

Although many experts in the field [17, 18] consider education and prevention initiatives is the most effective way to decrease mortality, it is still necessary to develop a tuned and dynamic model to keep track of and overcome statistical obsolescence.

From a Data Mining (DM) perspective, risk can be thought as a derivation of a set of variables heuristically selected as the best describing accident origin. When this is properly studied it is possible to predict not only a disaster but also its characteristics [19–21].

Many approaches in DM are used to predict events and find out its current and/or subsequent facts, like in [22, 23], etc.

Instead of that, this paper combines two reasoning systems: Expert Systems (ES) and Harmonics Systems (HS). The first one derives from the well-known technology started in 1980s but the second is quite a new technology presented in [24].

While Expert Systems (ES) remains focused on explicit rules of expert knowledge related to statistical prediction of the risk, the Harmonics System (HS) takes a heuristic data-driven approach. In HS, the information of interest is not the complex data produced along the development of an event, but only its timing patterns as a consequence of deep variables relationship during the process of an accident. This perspective is here performed by Harmonics Systems (HS), using combinations of variables (selected by an Expert System) as patterns. HS is a type of mining focused on rhythm, accelerations, static periods, and others aspects related to time features of selected patterns. HS also allow real-time processing, which is well fitted for applications that require prompt answers upon data collecting (that is the case of a driver o pedestrian collecting environmental data during displacement from one point to another). It is also included preliminary statistical results from real cases taken from [25]. Taking an Expert Systems in combination with HS [1], the traditional risk knowledge from the typical problems can be enhanced with dynamical timing patterns derived from previous activity and its variables. This kind of plastic, flexible, and self-trained learning model may serve from data. A resonance in this context can be thought of as a pattern matching with weighted features and chaining patterns. This modeling approach is being applied as the KRONOS prototype, to evaluate pedestrian and car risk. As a prototype is partially implemented, statical evaluation does not fully include HS add-ons or Fuzzy Logic at the pattern's matching process.

In the following, we shall present the basics for ES (Sect. 2), Harmonic systems (Sect. 3), the global architecture of KRONOS prototype that implements it as the core of its Expert System (Sect. 4) and a test application with real data (Sect. 5) followed by Conclusions and future work.

2 Expert System (ES)

This section presents a summary of ES as used in this project and its main characteristics.

2.1 ES Goal

The ES is used to reflect long-term expert knowledge and reasoning. The rules in KRONOS are defined using the variables and knowledge presented in [26]. Table 1 shows a reduced set of rules.

It is important to note that rules are not probabilistic or fuzzy. They fire just using the traditional approach.

Table 1 Some of the rules in the ES

IF	THEN
(individual.alcohol < 0.15) and (car.time > weather.sunset)	*car.risk.level = LOW* *individual.alcohol.level = LOW* *individual.risk.description = "You're perfects conditions"* *individual.risk.type = 0*
(individual.alcohol > 0.15) and (car.time > weather.sunset)	*car.risk.level = LOW* *individual.alcohol.level = LOW* *individual.risk.description = "You experimented a decreased reflexes"* *individual.risk.type = 1*
(individual.alcohol > 0.20) and (car.time > weather.sunset)	*car.alcohol.level = LOW* *individual.alcohol.level = LOW* *individual.alcohol.description = "Decreased reflexes, dysmetria and velocity underestimation"* *car.alcohol.type = 1*

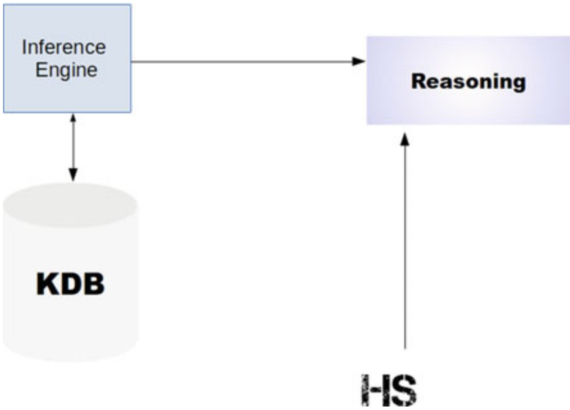

Fig. 1 ES architecture

2.2 *ES Architecture*

As any other ES, the prototype implements a core using the modules shown in Fig. 1.

3 Harmonic Systems (HS)

This section presents a summary of HS presented previously in [24, 26–28].

3.1 HS Goal

Once a problem has two characteristics: Real-time requirement and complex time behavior, HS can be applied. One reason is its lightweight algorithm and simple evaluation.

The other main reason is the production of meta-data generated as a part of model, reflecting change in time (as opposed to those models created through mining) while preserving other information related to the identity of the problem. When data results are processed it can be done in one of more patterns with the same or a different time variation.

An extra practical feature is the optional preprocessing with filters, to select specific time subsequence and ignore the rest of the data. This reduces significantly the amount of data being processed.

3.2 HS Problems

As the focus of HS is time and its change, it is suitable for problems that require a model of changes in time. Of course it demands one or more variables with specific patterns: co-occurrences, mutual exclusions, sequences, etc. One restriction in applicability is that variables must have a numerable finite data domain.

Taking into account the mentioned tips, HS can be used to test specific patterns of interest (for instance production failures, software/hardware faults, hang out of processes, deadlocks, etc.) while the main system works. As a consequence it can react to changes of behavior upon those patterns.

3.3 HS Functioning

Here there is a very short description of HS functioning. Since HS is out of the scope of this paper, readers interested in details may find them in [24].

Let a problem $\boldsymbol{R}$ consisting of a set of variables $\{v_i\}$ each one with a specific numerable finite data domain $\boldsymbol{D}_i$.

Let any relevant event $\boldsymbol{e}$ represented by one o more patterns $\boldsymbol{M}_j$, each one a combination of any subset of $\{v_i\}$ with specific values $\{v'_i\}$:$v_i \in \boldsymbol{D}_i$.

Then, the HS model to approximate consists of the union $\cup_i\{\boldsymbol{M}_i\}$, for all the events $\boldsymbol{j}$ whose patterns are being analyzed. And $\boldsymbol{M}_i$ defined as

$$U_{\mathbf{k}} = U_{\mathbf{k}} + \boldsymbol{\eta}_{\mathbf{u}}\left[U_{\mathbf{k}} - \boldsymbol{\Pi}_{\mathbf{l}} P_{\mathbf{0}}(t_{\mathbf{l}} | M_{\mathbf{i}})\right]$$

Table 2 Pattern 1

T	Property-1	Property-1
t1 = λ1	PROC = A	USR = 034
t2 = λ2	PROC = C	USR = 035
t3 = λ3	PROC = A	USR = 035

Where $\boldsymbol{U}_{\mathbf{k}}$ stands for time behavior model for pattern $\boldsymbol{M}_{\mathrm{i}}$, $\boldsymbol{t}_{\mathrm{l}}$ is the time elapsed from a previous occurrence of patterns feature $\boldsymbol{l}$, $\boldsymbol{\eta}_{\mathrm{u}}$ is a elasticity parameter for model $\boldsymbol{U}_{\mathrm{k}}$, $\boldsymbol{P}_0(\boldsymbol{t}_{\mathrm{l}}|\boldsymbol{M}_{\mathrm{i}})$ is the Poisson distribution probability for $(\boldsymbol{t}_{\mathrm{l}}|\boldsymbol{M}_{\mathrm{i}})$.

A set of additional parameters for $\boldsymbol{M}_{\mathrm{i}}$ are $\{\lambda_{\mathrm{l}}\}$ where

$$\boldsymbol{\lambda}_l = \boldsymbol{\lambda}_l + \boldsymbol{\eta}(t_l - \boldsymbol{\lambda}_l)$$

with η being a global parameter for all the model that represents a global adaptation coefficient for all the λ_{l} parameters.

In this context, an harmonic is the occurrence in time t_{l} of certain combination of properties that are of interest, and is referred to as pattern $\boldsymbol{M}_{\mathrm{i}}$. For example a pattern may be the one represented in Table 2.

In the table, PROC is a variable representing a software process of a complex system, A is a specific procedure, USR is an user ID that is being running that procedure, and t_1, t_2, t_3 are the typical time elapsed between them (in this example they are set of λ_1, λ_2, λ_3, respectively, as an initialization procedure). Another point of view of this problem is to model the sequence t_1, t_2, t_3, occurrence and variations. It may be used for instance to trigger actions while the sequence is happening or after it. When events match the pattern (a harmonic is found) there is a resonance, and the model may learn any variation in critical parameters.

A resonance has the following steps:

- Pattern detection: Patterns are evaluated against current data (In example 1: Property-1 = A, Property-2 = 034), compare the probability of the pattern against its threshold U (0.3 for example).
- Resonance: when there is resonance, the model parameters are updated.
- Fire an activity (optional) to produce meta-data and tracking data.
- Time information is processed (t_1, t_2, and t_3 in the example) as time-stamp shifts of the events.
- The size n is compared against a certain cut-off threshold nc (i.e., nc = 80). When n < nc, small (n) is true, otherwise it is false. When small (n) gives true, the Binomial dispersion of harmonics is assumed, otherwise it is considered to be Poisson.

3.4 HS Combined with FL

Harmonics may be implemented as data vectors with a predefined timing. But those times are the leading factor for the pattern to be in resonance or not. Thus, whenever time may relax and the relevant feature of the pattern is the set and organization of variables, then time may be considered not as sharp, but as a fuzzy number.

This Fuzzy Harmonic System (HFS) may be considered as a new approach, and takes traditional Fuzzy Logic (FL) as a shortcut to improve model stability when the patterns have many fluctuations in time. That way, for a narrow set of problems, it is possible to define a self-tuning set of λ_i parameters that converge asymptotically to a static value.

The reason to consider FL is historical. FL is usually considered an extension of classical logic. It can also be thought from the set theory as a sharp set with a fuzzy boundary. In NLP it is usually applied to model semantics and subjective information [29]. Computational Intelligence usually applies FL to a variety of problems, usually with complex and imprecise values.

Among others, additional benefits of fuzzy logic are its simplicity and its flexibility. The main reason to choose Fuzzy logic is not its ability to handle incomplete data, but the possibility to undertake problems with imprecise data, to model nonlinear functions of arbitrary complexity.

Fuzzy logic models are usually called fuzzy inference systems. They consist of a number of conditional "if-then" rules. For the designer who understands the system, these rules are easy to write, and as many rules as necessary can be supplied to describe the system adequately.

The main characteristic in fuzzy logic, unlike standard conditional logic, is that the truth of any statement has a degree. The conditions are usually coded as inference rules of the form A - > B (A implies B). But in FL, it can be said as (0.2 * A) - > (0.5 * B).

For example: the rule

$$\mathbf{A - > B}$$

with

A: module A takes 34 Mb
B: the weather daemon is on

can be restated as

if ***(module A takes 34 Mb)***
then (***the weather daemon must be started*****)**

Here are two variables: memory consumption for module A, and weather daemon status.

Both can relate to ranges of values (the first in Megabytes and the second a set of possible status).

Fuzzy inference systems rely on membership functions that represents to the computer how to calculate the correct value, between 0 and 1. It is often said that the degree to which any fuzzy statement is true is the denoted by a value between 0 and 1.

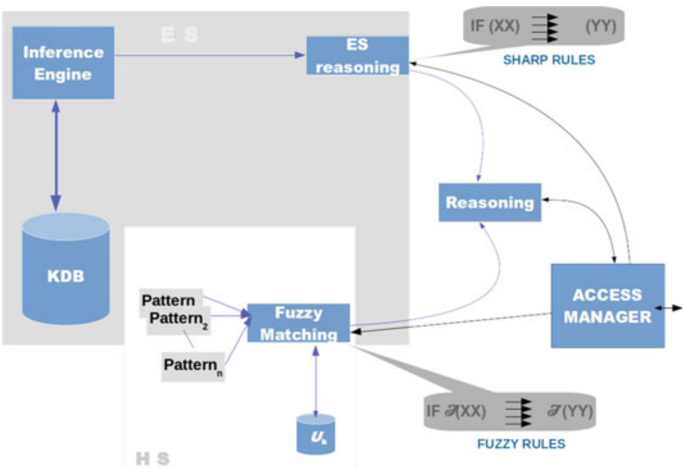

Fig. 2 ES architecture with Module Fuzzy Logic

Another perspective for the same approach is the Fuzzy Set Theory. It was formalized by Professor Lofti Zadeh at the University of California (1965). Zadeh proposed this paradigm with successful application worldwide. In this case, there is a set of rules and regulations that define boundaries and depict the best solution to solve problems restricted to those boundaries. The use of Fuzzy Set Theory from conventional bivalent, sharp sets theory is also considered a paradigm shift.

The use of FL will allow the system to pay attention to boundary events and contexts that have complex resolution. This extra flexibility will pay the cost known as the Non-Free-Lunch theorem [30]. Although there are many speculations regarding it, in general sense it describes that the best an algorithm solves a type of problems, the worse it performs in general. Taking that into account, FL processes facts in a proper way to overcome transient alterations in the system [31, 32].

Figure 2 shows the architecture of the FHS proposal. It has the previous ES with the same components, coordinated with HS. As earlier, both receive the input information from an Access Manager (a module to interface a global controller focused on compatibility with external devices and systems.

The key difference is the set of furry rules that are now biasing the pattern matching process inside the HS. Thus the entire process' performance is being altered.

3.5 *HS Specific Features*

Harmonics may be implemented as data vectors with a predefined threshold of tolerance for diverges. But some other characteristics of this approach are

- There is no precise time but relative: Time is the duration of certain event, opposed to classic techniques [33] where the value of a certain property is compared at time t_i respective to t_{i-1}, and the magnitude of a property associated. As a consequence there is no comparison between length series or corrections in them due to a different length. Therefore there is no normalization.
- No corrections required: Since no component alignment is required between patterns, distance has no need for corrective techniques such as dynamic time warping, longest common subsequence similarity, local scaling functions, global scaling function, etc.
- Flexibility: HS manages properties being measured as a pattern, which identifies the components in a time series, and models the time dispersion instead of the set of properties inside the pattern.

KRONOS models patterns' time features, and could be analogous to probabilistic similarity measure where methods are model based (they can incorporate prior knowledge into the similarity measure). However, it is not clear whether other problems such as a time series indexing, information retrieval, and clustering can perform efficiently. They use a general similarity approach involving a transformation rules language [1], and hundreds of algorithms from DM to classify, cluster, segment, and index time series.

3.6 HS Filters

Certain problems have too much information throughput, generating an extensive dataset, and making it very hard to perform efficient analysis. In these cases the model may be extended to go through one or more data and/or pattern filters. The effect of this preliminary step is biasing information to focus on specific harmonics. There are three types of filters [24]:

- High-pass filters: They leave the data that are beyond a certain distance (δ) that ($t_{i+\delta}$ < tactual|pattern). Since the pattern's property $p_1..p_i$ is met, t_i exceeds the model value.
- Low-pass filters: They leave the data that are closer than a certain distance ($t_{i-\delta}$ > tactual|pattern). Since the pattern's property $p_1..p_i$ is met, t_i is lower than the model value.
- Band-pass filters: They leave the data that are within a certain distance range ($t_{i+\delta}$ > tactual|pattern > $t_{i-\delta}$). Since the pattern's property $p_1..p_i$ is met, t_i is within the model value with a certain distance.

4 The KRONOS Prototype

This section presents a summary of KRONOS presented as a proposal in [26–28].

4.1 Architecture

Kronos is a prototype that implements a model for time predictions. After collecting data from many sources, an Expert System interacts with a Knowledge Base and an intelligent HS subsystem. Its goal is to evaluate any traffic and pedestrian risks. The global design is able to interact with diverse and mobile devices, other information systems, user data, and Internet (Fig. 3).

Main components are

- **Web**: It is a source of information and requests. A web server, web service, a local server, or other host may connect with the prototype using a proper interface represented in the picture with this module.
- **Host**: The prototype has a rule-based Expert System for data prediction [24], to evaluate risks based on expert knowledge, inputs and historical statistics. It interacts with the HS subsystem to dynamically build a more precise model.
- **Input Device**: Information regarding current position, status, and requests may be provided to the prototype by one or more mobile devices, sensors, etc. Each one requires a specific interface.
- **Output Device**: As it may be used by pedestrians and drivers, it is expected to output information upon requests though mobile devices' interfaces.
- **External System**: Already existent systems may interact with the prototype using the interface represented here as this module.

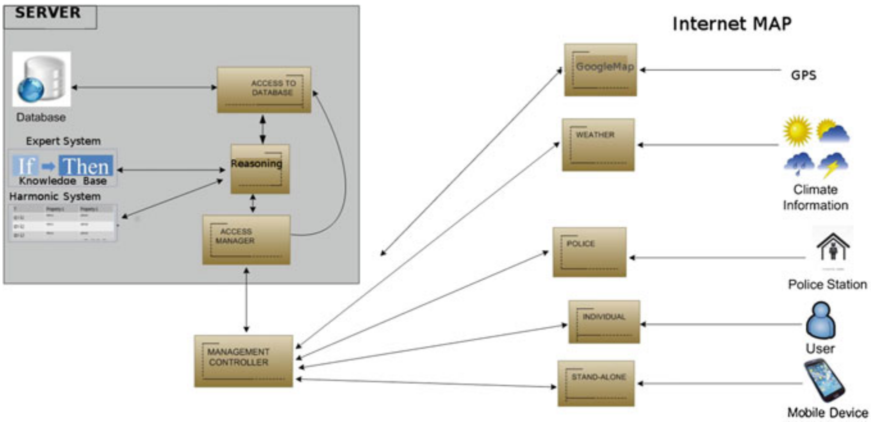

Fig. 3 KRONOS architecture

- **DBMS**: it is a Postgres database with a set of store procedures and triggers that automatically check and transform data.

4.2 Characteristics of Time Mining

Data may be collected from many sources and converted to be compatible with internal and DBM's requirements. They have the following characteristics:

- Belong to an external entity.
- Have a time stamp with a least date, time, minutes, and seconds.
- May be associated with a cyclic and complex event producing a measurable set of features.
- The source of the data is one or more identifiable and distinguishable sources.
- Events are well defined and limited in time.
- Their duration is variable or constant but they start and end at a defined point in time.
- They may undergo duration and frequency changes, but not changes in identifying characteristics.

5 Test and Evaluation

As mentioned previously, the prototype partially implements the model. The ES core is functional, it has many rules according to expert recommendations. The interfaces to user, DBMS, Maps, DBMS, are working and also part of the Harmonics system. The tests in this section use ES knowledge but not the Harmonics system, since the goal of this paper is to provide the improvement acquired by considering GPS in the risk inference.

5.1 Database

The test set are risk situations evaluated post-Morten to be able to find the accuracy of the results.

All the dataset belongs to real situations of traffic in Concepción del Uruguay city (Entre Ríos, Argentina). Figure 4 shows the map of the city.

As can be seen the map has a grid. Each cell represents a zone, described in the Database as follows: (ID-Zone, Description, Latitude, Longitude, Risk). ID-Zone is an integer that identifies the cell. Description outlines the zone. Table 3 describes the reference values and the labels in Fig. 4.

Fig. 4 City map

Table 3 Zone labeling

Risk	Risk Type	Label
0	No	N
1	High	A
2	High	A
3	Medium	M
4	Low	None
5	Low	None

The testing resulting in the results summarized in Table 4.

In Table 4, column Desc. (Description field), has the values: (c)enter, (r)oundabout, c(o)untry road, country (z)one. Column S (Subject) has the values: (d)river, (p)edestrian.

To assess the impact of considering GMT (Zone) information, the dataset was reevaluated avoiding that variable from the rules. Then, the predictions were compared to expert predictions. The number of hits and errors are in Table 5.

The legend (~G) means without GMT information, and (G) means with GMT information. Results indicate higher accuracy (94% vs. 42%). Also the system trends to underestimate risk. Analyzing these three test cases, the deviation occurs for drivers that do not use helmet/belt and are in a safe zone of the city, during the daylight hours but when weather has reduced visibility. Figure 5 shows results test by test.

The (E)xpert prediction is the darkest curve. Without GMT information (~ G) the ES prediction results are more erratic.

Table 4 Testing results

Risk P/G	Dd/mm/yy	Time	Weather	Desc.	S	km/h	Belt/helmet	Alc. (g/l)
0	01/05/15	13:55	Sunny	C	D	40	Si	0.9
0	22/05/15	12:10	Sunny	R	D	90	No	0
0	12/06/15	22:29	Sunny	O	D	40	No	0.3
0	24/11/15	08:32	Foggy	O	D	30	No	0
0	01/05/15	13:56	Sunny	C	P			0.3
0	08/05/15	11:30	Cloudy	Z	P			1
1	12/06/15	22:29	Cloudy	Z	D	122	No	0
1	28/06/15	17:43	Cloudy	C	D	140	No	0.4
1	12/06/15	23:44	Cloudy	C	P			0.2
1	12/06/15	23:45	Cloudy	C	P			0.5
1	23/11/15	23:32	Cloudy	C	D	190	No	0
1	23/11/15	23:50	Cloudy	C	D	230	No	0.9
1	12/06/15	22:20	Cloudy	Z	D	130	No	0.5
3	12/06/15	23:03	Cloudy	C	D	100	No	0
3	12/06/15	23:13	Cloudy	Z	D	100	No	0.2
3	08/09/15	21:45	Cloudy	C	D	150	No	0.1
3	02/10/15	19:21	Sunny	C	D	190	No	0.9
3	02/10/15	19:48	Sunny	C	D	190	No	0.9
3	12/06/15	23:44	Cloudy	C	P			0.51
3	12/06/15	23:44	Cloudy	C	P			1.25
0	25/11/15	11:41	cloudy	C	P			0
0	25/11/15	12:36	Fog	Z	P			2
0	25/11/15	19:36	Sunny	C	P			1.9
0	25/11/15	02:36	Sunny	C	P			1.9
0	25/11/15	02:36	Sunny	Z	P			1.5
1	25/11/13	13:53	Sunny	C	D	120		1
1	12/08/15	19:30	Cloudy	Z	D	140		0.19
1	17/08/10	19:30	Cloudy	Z	D	140		0.16
1	17/08/10	19:30	Cloudy	Z	D	125		0.18
1	17/08/10	20:30	Cloudy	Z	D	125		0.25
2	25/11/15	20:14	Cloudy	C	D	90		1
2	25/11/15	14:21	Cloudy	C	D	40		2
2	25/11/15	14:21	Cloudy	C	D	50		2
2	25/11/15	14:21	Cloudy	C	D	50		2
2	25/11/15	14:21	Cloudy	C	D	60		2
3	25/11/15	14:36	Sunny	Z	D	40		0.1
3	25/11/15	20:36	Sunny	Z	D	40		0.1
3	25/11/15	20:36	Sunny	Z	D	60		0.1
3	25/11/15	20:36	Sunny	Z	D	90		0.1

(continued)

Table 4 (continued)

Risk P/G	Dd/mm/yy	Time	Weather	Desc.	S	km/h	Belt/helmet	Alc. (g/l)
3	25/11/15	14:36	Sunny	Z	D	35		1
5	31/05/11	23:22	Cloudy	C	P			0.13
5	04/05/15	23:22	Cloudy	C	P			0.1
5	25/11/15	15:24	Sunny	C	P			0.1
5	25/11/15	20:00	Sunny	C	P			0.14
5	25/11/15	20:00	Sunny	C	P			0.1
5	09/07/14	23:48	Sunny	C	P			0.11
5	09/07/14	23:48	Sunny	C	P			0.1
5	09/07/14	23:48	Sunny	C	P			0.05
5	09/07/14	23:48	Sunny	C	P			0.03
5	09/07/14	23:48	Sunny	C	P			0.09

Table 5 Results with and without GMT

	Error (G)	%	Error (G)	%
Overestimated	11	22.00	0	00.00
Underestimated	18	36.00	3	06.00
OK	21	42.00	47	94.00

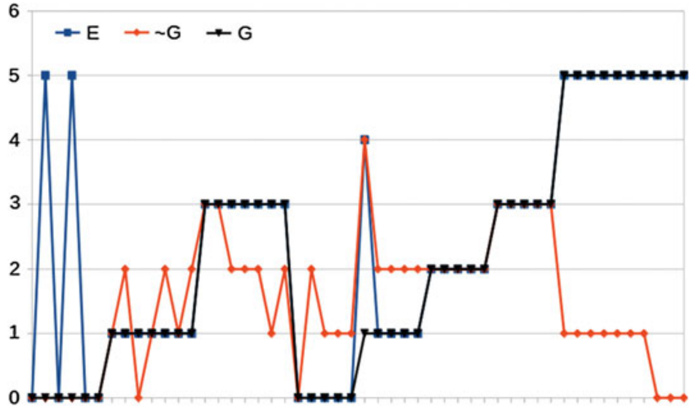

Fig. 5 Risk according Expert (E), ES without GMT (~G) and with GMT (G)

6 Conclusion and Future Work

This paper presents an outline of the KRONOS project, a prototype of a model with a dual risk evaluation mainly using statistical and heuristic approaches respectively. The key features of each one were presented as well as basic statistical information regarding how the statistical inferences can be improved using GPS information.

The comparison between the accuracy acquired in previous work has increased to 52%, and it is possible to say that GPS information has a good impact in the results. This trend must be verified with a larger number of test cases.

As a future work it remains to test HS as it was performed with ES, and find out how both perspectives may be combined to provide better results. Also a FL treatment for patterns is pending for implementation and statistical evaluation.

References

1. López De Luise, D.: MLW and bilingualism. Advances Research and Trends in New Technologies, Software, Human-Computer Interaction, and Communicability. IGI Global, USA (2013)
2. Wash, O.: Assessing pedestrian risk locations: a case study of WSDOT efforts. Department of Transportation. Washington State Library. Electronic State Publications (1998)
3. Organization for Economic Co-operation and Development (OECD): http://www.oecd.org (2014)
4. Alex Quistberg, D., Jaime Miranda, J.: Beth Ebel. Reducing pedestrian deaths and injuries due to road traffic injuries in Peru: interventions that can work. Revista Peruana de Medicina Experimental y Salud Pública. Rev Peru Med Exp Salud Publica vol. 27 n. 2 Lima Apr/Jun 2010. ISSN 1726-4634
5. Oxley, J.: Improving Pedestrian Safety. Curtin—Monash Accident Research Centre. Fact Sheet No. 6. (2004)
6. Cathy, T., D. Packman, Risk and safety on the roads: the older pedestrian. Foundation for Safety Research. New Castle University (1995)
7. Generic Risk Assessment 4.1. TSO Publisher: www.tsoshop.com.uk (2009)
8. Antov, D., Rõivas, T., Antso, I., Sürje, P.: A method for pedestrian crossing risk assessment. Transaction of Wessex Institute. doi:10.2495/UT110501 (2011)
9. Traffic Management Hazard Identification & Risk Assessment Control Form: Swinburne University of Technology. http://www.docstoc.com/docs/24507714/Traffic-Management-Health–Safety-Checklist (2009)
10. Alavi, H., Charlton, J., Newstead, S., Archer, J.: A Pedestrian Data System for Safety Analyses. Monash University Accident Research Center (MUARC), Melbourne, Victoria, Australia (2014)
11. Jiao, J., Moudon, A.: Using a Case-control approach and GIS Methods to Assess the Risk of Pedestrian Collision In Seattle, USA
12. Kim, K., Sul, J.: Development of Intersection Traffic Accident Risk Assessment Model Transportation & Environment Research Institute Ltd. kijoonkim@hotmail.com, 82-(0) 2-10-8752-1851 (2001)
13. Hautzinger, H.: Analysis Methods for Accident and Injury Risk Studies. Project No. 027763 TRACE. Deliverable 7.3.2007
14. Journey Risk Management® (JRM®): http://www.irte.com/journey-risk-management.html
15. Rodríguez-Hernández, M., Campuzano-Rincón, J.: Primary prevention measures for controlling pedestrian injuries and deaths and improving road safety. Revista Salud Pública. Rev. salud pública vol. 12 no. 3 (2010)
16. Thomas, L., Hamlett, C., Hunter, W., Gelinne, D.: Final Report to North Carolina Department of Transportation. North Carolina Department of Transportation, Traffic Engineering and Safety Systems Branch (2009)
17. Hart, J.: Measuring Pedestrian Risk and Identifying Methods to Prevent Pedestrian Accidents in Langley Park. National Fire Academy (2004)

18. Health and Safety Comission: Reducing at-work road traffic incidents. Report to Government and the Health and Safety Commission. DTLR (2001)
19. Han, J., Kamber, M.: Mining Stream. Time-Series, and Sequence Data. In Data Mining, 2nd edn. Concepts and Techniques, 2nd edn (2011)
20. Shieh, J., Keogh, E.: iSAX: indexing and mining terabyte sized time series. In: Proceedings KDD'08, pp. 623–631. ACM (2008)
21. Tak-chung, F.: Engineering applications of artificial intelligence. Eng. Appl. Artif. Intell. **24**, 164–181 (2011)
22. Bollogás, B., Das, G., Gunopulos, D., Mannila, H.: Time series similarity problems and well-separed geometric sets. Nordic J. Comput., 409–423 (2003)
23. Jagadish, H., Mendelzon, A.: Similarity based queries. In: Proceedings of the 14th PODS, vol. 95, pp. 36–45 (1995)
24. López De Luise, D.: Harmonics systems for time mining. Int. J. Modern Eng. Res. (IJMER) 3 (6, 3), 2719–2727 (2013). ISSN:2249-6645
25. Instituto Nacional de Estadísticas y sensos: México. http://www3.inegi.org.mx
26. Acuña, I., García, E., López De Luise, D., Paredes, C., Celayeta, A., Sandillú, M., Bel, W.: Traffic & Pedestrian risk inference using Harmonic Systems. SOFA, Romania (2014)
27. Celayeta, A., Paredes, C., López De Luise, D., Bel, W.: Traffic and Pedestrian risk evaluation with Harmonic Systems (Cálculo de Riesgo en tráfico y Peatón usando Sistemas Armónicos). ARGENCON (2014)
28. Mansilla, D., Sandillú, M., López De Luise, D., Bel, W.: Un Modelo con Conocimiento Experto y Sistemas Armónicos para Evaluación de Riesgos de Tráfico. EnIDI Argentina (2015)
29. Zadeh, L.A.: Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. Syst. Man Cybern. **3**(1), 28–44 (1973)
30. Wolper, D.H.: The supervised learning no-free-lunch theorems. In: Proceedings of the 6th Online World Conference on Soft Computing in Industrial Applications (2001)
31. Zadeh, L.A.: A note on prototype set theory and fuzzy sets. Cognition **12**, 291–297 (1982)
32. Zadeh, L.A.: Fuzzy Sets and Applications (Selected Papers, ed. by R.R. Yager, S. Ovchinnikov, R.M. Tong, H.T. Nguyen. Wiley, Nueva York (1987)
33. Ratanamahatana, C.A., Lin, J., Gunopulos, D., Keogh, E.: Mining time series data. In: Data Mining and Knowledge Discovery Handbook, pp. 1049–1077 (2010)
34. De Nicolao, G., Ferrara, A., Giacomini, L.: On-board sensor-based collision risk assessment to improve pedestrians' safety. IEEE Trans. Veh. Technol. **56**(5), 2405–2413 (2007). doi:10.1109/TVT.2007.899209