Identifying Outlying Power Two Wheeler Riding Behaviors at the Emergence of an Incident

Eleni I. Vlahogianni*
National Technical University of Athens,
5 Iroon Polytechniou Str, Zografou Campus, 157 73 Athens Greece
Tel: +30 210 772 1369; Fax: +30 210 772 1454
e-mail: elenivl@central.ntua.gr

George Yannis
National Technical University of Athens,
5 Iroon PolytechniouStr, Zografou Campus, 157 73 Athens Greece
Tel: +30 210 772 1326; Fax: +30 210 772 1454
e-mail: geyannis@central.ntua.gr

John C. Golias
National Technical University of Athens,
5 Iroon Polytechniou Str, Zografou Campus, 157 73 Athens Greece
Tel: +30 210 772 1276; Fax: +30 210 772 1454
e-mail: igolias@central.ntua.gr

N. Eliou, Ph.D.
University of Thessaly,
PedionAreos 38334 Volos, Greece
neliou@uth.gr

Words: 5007
Tables: 5
Figures: 2

**Corresponding author*

ABSTRACT
High resolution riding data from a large-scale Power Two Wheeler's (PTW) naturalistic driving experiment are exploited in order to identify critical riding patterns emerging at the beginning and during an incident. A two-step analysis is adopted: first, a clustering approach is undertaken in order to reveal the critical rider's actions at the beginning and during an incident. Second, the revealed actions are associated to specific riding situations in order to identify the critical riding patterns. Both methodological steps are modeled using Bayesian Networks. Results reveal three different prevailing riding actions for describing the onset of an incident and an equal number of actions that a rider executes during the course of an incident to avoid a crash. Furthermore, the proposed methodology efficiently relates the observed sets of actions with the different riding incidents and produce riding patterns (moving or stationary obstacle, overtaking and opposing traffic) that are characterized by different initial actions, as well as by different rider's action likelihood during the incident.

Keywords: naturalistic driving, road incidents, riding behavior, power two wheelers, cluster analysis, Bayesian networks

INTRODUCTION

Risky rider's behavior has been for long a critical consideration in Power Two Wheelers' (PTW) safety. The majority of research attempts focus on demystifying the riding behavior in critical riding situations in order extract useful knowledge on the manner PTW users react to internal of external stimuli and develop efficient countermeasures for PTW safety; attempts range from macroscopic and in-depth analyses of PTW accident datasets based on police reports or queries *(1-6)* to advanced riding simulator experiments that target the continuous and microscopic monitoring of riders' behavior (7-*10)*. The limitation of police reports and queries is that riders/drivers involved in accidents usually – intentionally or unintentionally – provide police with misleading information, as well as little accurate information on the pre-crash conditions *(11)*. As for simulators, although they provide a setting for studying in detail the reaction of riders during extreme situations, they hardly can claim similarities to reality, especially with regards to the dynamics of PTWs that are very difficult to replicate *(10)*.

Recently, PTW riding studies have been supported by significant technological advances that enable the efficient, least intrusive and continuous monitoring and recording of data on the manner a rider behavior on its own physical environment, the road. Experiments of such a technological depth known as naturalistic embed high resolution video recorders, sensors and data storage units that may efficiently monitor and record every activity of the rider, such as brake activation, steering, speed, acceleration, yaw and so on. Although some naturalistic experiments have been conducted in Europe, the US and Australia to understand driver behavior *(12-14)*, so far no quantitative results are publicly available concerning rider behavior and especially regarding the prevailing riding behaviors observed at the occurrence of an incident *(12)*,*(15)*,*(16)*. This information may be considered as critical in order to demystify the manner a rider reacts to and internal or external stimulus.

In all relevant approaches documented so far, some shortcomings regarding the detection and identification of reactions of PTW riders can be found. All studies implement typical fixed driving/riding parameters' thresholds – regardless of the type of rider and the type of area or other roadway of rider characteristics- and, based on those values, the incidents are extracted and further analyzed *(11)*,*(17)*. This technique lacks consistency with the fact that each driver/rider has its personal stock of values, ideas, beliefs and practices, reflecting rigorously on its behavior on the road, such as the braking, overtaking and so on, that may not converge to a "typical rider's behavior." This leads to a different definition of the notion "incident" for each driver. Furthermore, the use of fixed thresholds to identify critical incidents severely biases the riding behavior that may be associated with the occurrence of a specific incident. Recently, a statistical approach to detect critical incidents from a multivariate set of riding parameters that define the riding characteristics of a specific rider has been proposed *(18)*; "incidents" are defined as those situations that the rider's actions deviate from his mean riding behavior. The mean behavior and its deviation are defined in relation to changes of the braking, wheel speed, steering and throttle. Although, causalities could not be established, this methodology is a first attempt to automatically detect irregularities and critical incidents from a vast amount of complex and high resolution naturalistic riding data.

Moreover, until now little is known on the manner a rider reacts to the emergence of a critical incident. More specifically, there is no knowledge coming from data collected on the interrelations between braking, speeding and maneuvering under different riding or roadway conditions. The present paper proposes a methodology for identifying the riding behaviors that arise at the emergence of a critical incident based on high resolution monitored riding data (100Hz) from a naturalistic driving experiment, consisting of information on wheel speed, acceleration, throttle, steering, braking and so on. The proposed approach is based on Bayesian Networks and aims at clustering the critical incidents' characteristics in order to reveal typical riding behaviors in the emergence and during a critical incident.

METHODOLOGY

A fundamental research question related to identifying criticalities in riding behavior is whether there is a way to relate critical riding actions and relate them to specific riding situation (incidents). In this paper, this question is treated in two steps. First, it is assumed that, at the emergence of an incident, the rider performs a far from typical (mean) riding action that is followed by a set of sequential actions during the incident in order to avoid crash. For this step, a clustering approach is undertaken in order to reveal the critical rider's actions at the beginning and during the occurrence of an incident. In the second step, the revealed actions are further associated to specific riding situations, for example overtaking, avoiding stationary obstacle and so on, taking into consideration the uncertainties arisen form the manner the rider will react to each situation.

Both methodological steps are modeled using Bayesian Networks (BNs). A Bayesian network $BN = \langle K, L, \mathbf{\Theta} \rangle$ is a directed acyclic graph $\langle K, L \rangle$ of $k \in K$ nodes that represent the x_i random variables of the network *(19)*. Nodes are connected by links $l \in L$ that describe the probabilistic relationship between interconnected nodes; this relationship is quantified using a conditional probability distribution $\theta_i \in \mathbf{\Theta}$ for each node k_i (Friedman et al. 1997): $\theta_{x_i|\Pi_{x_i}} = P_B(x_i|\Pi_{x_i})$, where $\Pi_{x_i} \in \Pi_{X_i}$, where Π_{X_i} stands for the set of parents of X_i in the network. Independency between variables is denoted by the lack of a link. A BN defines a unique joint probability distribution over *X* given by *(20)*:

$$P_B(x_1,...,x_n) = \prod_{i=1}^{n} P_B\left(x_i \middle| \Pi_{x_i}\right) = \prod_{i=1}^{n} \theta_{i|\Pi_i}\left(x_i \middle| \Pi_{x_i}\right) \quad (1)$$

BNs are powerful in handling incomplete data and uncertain phenomena. Moreover, due to their probabilistic nature, they can easily integrate both qualitative information and quantitative information in modeling. BNs have been successfully applied to traffic analysis and forecasting *(21-22)* incident detection *(23)*.

The BN can act as a classifier; given the characteristics $x_i \in \mathbf{X}$ as inputs (for example the riding parameters) and a set of classes *Z* (for example the riding situations), a new unclassified observation *S* can be assigned to a class by the rule *(20)*:

$$classify(x_1,...,x_n) = \arg\max_n \, p(z)\prod_{i=1}^{n} p\left(x_i|z\right) \quad (2)$$

A BN can be also used as a clustering model. Clustering is a task to partition the objects in the dataset *D* into clusters of similar objects. By using BN, each object with attributes x may be classified to its most probable cluster (class) k^*, based on the estimated parameters θ_i, by using a membership probability as a score (unsupervised classification):

$$k^* = \underset{1\le k\le K}{\arg\max}\, p\left(k \middle| \mathbf{x}, \hat{\boldsymbol{\theta}}\right) = \underset{1\le k\le K}{\arg\max}\, p\left(k, \mathbf{x} \middle| \hat{\boldsymbol{\theta}}\right) \tag{3}$$

Parameters are learnt via an Expectation-Maximization algorithm *(24)*.

The amount of information flow between two nodes x_i and x_j can be measured by mutual information. The mutual information $I\left(x_i, x_j\right)$ between variables x_i and x_j measures the expected information gained about x_j, after observing the value of the variable x_i *(20)*:

$$I\left(x_i, x_j\right) = \sum_{x_i \in X, x_j \in X} P\left(x_i, x_j\right) \log \frac{P\left(x_i, x_j\right)}{P\left(x_i\right) P\left(x_j\right)} \tag{4}$$

The mutual information between two nodes can tell us if the two nodes are dependent and if so, how close their relationship is. The information flow with respect to the set of "evidence" variables (condition-set), in this case the Class membership Z $I\left(x_i, x_j | z\right)$ is given by conditional mutual information *(20)*:

$$I\left(x_i, x_j | z\right) = \sum_{x_i \in X, x_j \in X, z \in Z} P\left(x_i, x_j, z\right) \log \frac{P\left(x_i, x_j | z\right)}{P\left(x_i | z\right) P\left(x_j | z\right)} \tag{5}$$

The learning procedure is based on quantifying the amount of information stored in each link of the network. First, an initial structure of the BN is developed and then the relationships are learnt. Structure evolves nodes dependencies and strength according to the mutual information *(19)*.

THE EXPERIMENT

Analyses are based on the data collected from a large scale naturalistic riding experiment conducted in the city and at the suburbs of Volos, a medium scale Greek city, during the period November 2010 - April 2011. The specific experiment is conducted using a BMW F650 Funduro. The available signals that are being stored are summarized in Table 1. Moreover, video installation is available and calibrated to capture the frontal environment (a minimum of 90° field of view) and the rider's face. The data acquisition of all signals is set to be recorded at an accuracy of 100Hz except for video signals which are sampled at 10Hz and GPS position which can be sampled at 1Hz (values indicate minimal requirements). Details on the experiment can be found in *(26)*.

Data selected for further study encompass trips of 20 minutes duration made during a period of three months by one rider in rural two-way roads. The specific route has shoulder width less than 1,2m, rolling terrain and high curvatures, mixed traffic conditions, not equally distributed traffic across the two directions, a significant number of uncontrolled access points and several zones where passing is permitted; this setting seems to be challenging for PTW riders in terms of riding patterns' complexity and difficulty in maneuvering. All trips were made in daylight with good visibility and fine weather conditions. The final dataset consists of 56 trips of 20 minutes duration, meaning a set of time series of riding parameters with $6.72{\cdot}10^6$ data. No crashes occurred during the experiment.

TABLE 1: The list of monitored variables and their description.

Variable	Description
Longitudinal acceleration (g) Lateral acceleration (g) Vertical acceleration (g)	Linear acceleration (three components)
Speed (km/h)	Longitudinal speed
Yaw rate (deg/s) Pitch rate (deg/s) Roll rate (deg/s)	Roll, yaw and pitch angles and rates
Throttle (%)	Throttle position
Brake Rear (%)	Brake pressure rear
Steering (%)	Steering angle
Brake activity Front (0/100)	Brake activation front
Brake activity Rear (0/100)	Brake activation rear
Wheel Speed (km/h)	Speed of the rear wheel

Based on a similar dataset, previous research has found that steering, throttle, brake activation and wheel speed are the optimum set of parameters in order to effectively detect the critical incidents, meaning situations with increased probability of leading to an accident, whereas the inclusion of traffic related variables have been found less influential in detecting incidents *(18)*. The detection of critical incidents was based on a robust outlier detection methodology based on the Mahalanobis distance d_i defined as (Barnett and Lewis 1994): $d_i = \sqrt{(x_i - \hat{\mu}) S^{-1} (x_i - \hat{\mu})}$, where, $X_i = (x_{i1},...,x_{ip}),\ i = 1,...,n$ is the multivariate space of p riding parameters that independently come from a multivariate normal distribution $X = N(\mu, \sigma^2)$, where μ is the mean and σ is the covariance matrix, $\hat{\mu}$ and S^{-1} are the sample mean and covariance matrix respectively *(25)*. The Mahalanobis distance can be approximated by an F-distribution$\left[p(n-l)(n+l)/n(n-p)\right]F_{p,n-p}$; at a significance level *α*, a determination as to whether a new observation X_i can be considered as outlier – critical incident - or not can be made based on the following formula: $d_i \leq \left[p(n-l)(n+l)/n(n-p)\right]F_{p,n-p}$.

For a given rider, the above algorithm may distinguish between typical – mean – riding patterns and irregular – far from the mean – riding behavior. An example of the riding incident detection methodology is seen in Figure 1 where the temporal evolution of the distance metric d_i is depicted; as can be observed, a large deviation from the mean signifies the onset of a critical incident. The larger the deviation the greater the change of the rider's riding style. Moreover, by this method, the duration of the event may be also measured as the time difference between the first deviation of distance metric from the mean value and the return to the mean riding pattern.

The specific detection methodology provides information in the form of joint consideration of steering, activation brake (front, rear), throttle and wheel speed for both the initial riding conditions at the occurrence of an incident, and the riding conditions during and incident. The initial conditions at the occurrence of an incident refer to a single action that is undertaken by the PTW rider as an immediate reaction to a stimulus, whereas the conditions during the evolution of an incident may encompass more than one action undertaken by the rider in order to prevent an accident. In the latter case the sequence of rider's actions may play an important role.

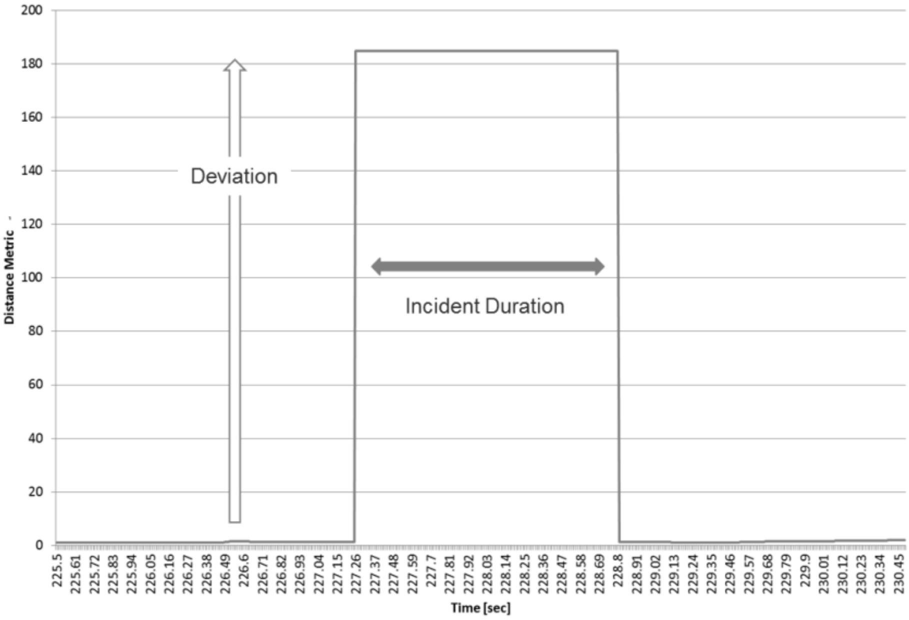

FIGURE 1: An example of a detected incident using the Mahalanobis distance metric that encompasses joint information on throttle, steering, braking activation and wheel speed.

Figure 2 depicts the distributions of all available data; as can be observed the distribution of the values of the available parameters may significantly vary between the beginning and during the observed incident.

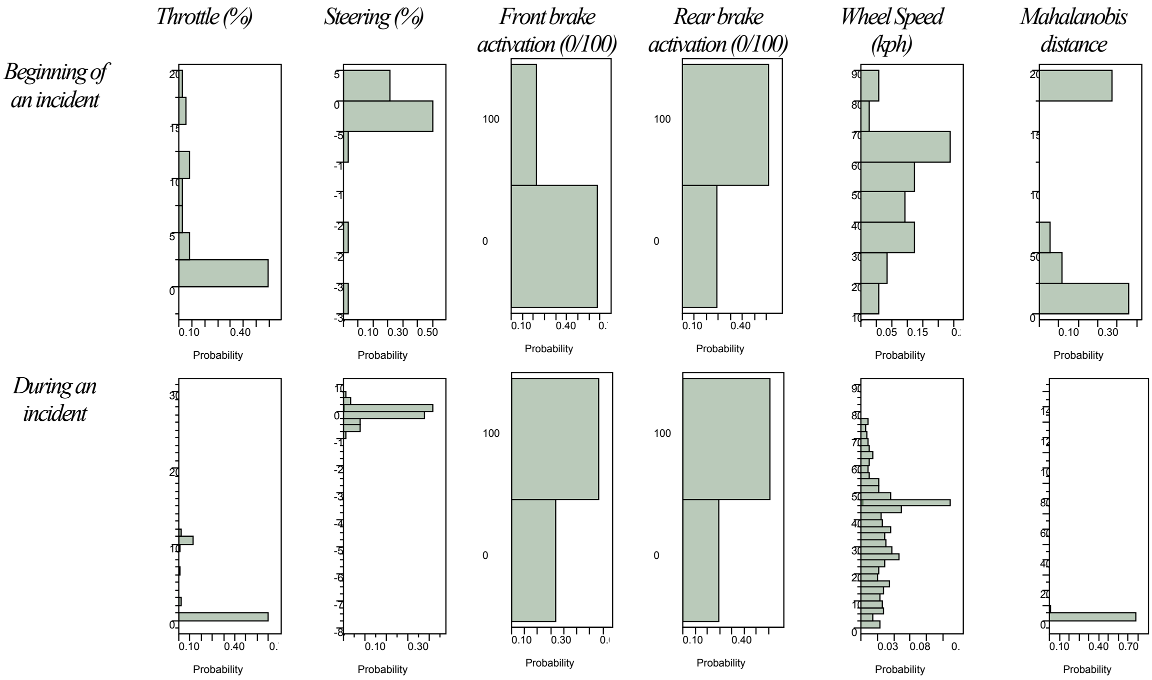

FIGURE 2: The distributions of the available variables for modeling.

IDENTIFYING CRITICAL RIDER ACTIONS

For each incident phase – beginning and during the incident - a BN clustering model is constructed in order to reveal groups of critical riding actions with respect to the riding characteristics defined by steering, braking, throttle and wheel speed. By critical, we mean the different actions that the rider engages at the beginning and during a critical incident. The trained BNs are able to produce clusters with 68% and 72% purity respectively; the relative high purity signifies that all clusters produced contain mainly cases from a sole class of actions and not from multiple classes of riding actions.

Results are summarized in Table 2 and Table 3 for the beginning and during an observed incident respectively. The profile of each cluster is established based on the binary relative significance, meaning the ratio between the mutual information brought by each variable and the greater mutual information; $\frac{\text{mutual information}_i}{\max\{\text{mutual information}_i\}}$, $i = 1, 2, ..., n$ where n is the number of variables describing riding. Moreover, for each influencing variable, its modal value, meaning the most probable value with respect to the response variable and its observed state; this modal value comes with its probability.

TABLE 2: Results from clustering for the initial conditions at the occurrence of an incident.

Node	Binary mutual information (%)	Binary relative significance	Modal Value[1]	
		S_1 (58.78%)		
Rear Brake Activation (0/100)	55.05%	1.00	100	100%
Front Brake Activation (0/100)	41.51%	0.75	0	100%
Throttle (%)	21.32%	0.39	<=2.47 (1/4)	81%
Steering (%)	15.81%	0.29	<=0.42 (3/4)	81%
Wheel Speed (kph)	4.35%	0.08	<=49.743 (2/4)	33%
		S_2 (21.56%)		
Front Brake Activation (0/100)	73.12%	1.00	100	100%
Wheel Speed (kph)	23.47%	0.32	<=33.45 (1/4)	57%
Throttle (%)	18.31%	0.25	<=2.47 (1/4)	100%
Rear Brake Activation (0/100)	7.61%	0.10	0	57%
Steering (%)	4.64%	0.07	<=0.42 (3/4)	57%
		S_3 (18.66%)		
Throttle (%)	79.25%	1.00	<=16.69 (3/4)	50%
Rear Brake Activation (0/100)	27.39%	0.34	0	83%
Steering (%)	14.78%	0.19	<=0.42 (3/4)	50%
Wheel Speed (kph)	11.45%	0.14	>65.91 (4/4)	50%
Front Brake Activation (0/100)	0.78%	0.01	0	67%

[1] *Values in the parentheses are the ranges of variable discretization the modal value belongs to.*

As seen in Table 2, the clustering revealed three distinct groups of riding actions associated with the prevailing conditions at the occurrence of an incident. In each group of actions, the independent variables are ranked in a different manner, indicating that variables' influence in the various groups of actions is different. The group S_1, encompassing the 59% of sample actions, reveals that the rider, at the beginning of an incident, activates the rear brake in medium speed, executing a minor maneuver. In this group, the most influential variable are the rear and front brake activation, followed by throttle and steering; wheel speed seems to be least influential.

TABLE 3: Results from clustering for the riding actions during a detected incident.

Node	Binary mutual information (%)	Binary relative significance	Modal Value[1]	
	C_1 (43.65%)			
Steering (%)	24.32%	1.00	<=1.84 (3/4)	86%
Rear Brake Activation (0/100)	21.53%	0.89	100	96%
Throttle (%)	15.13%	0.62	<=5.56 (1/2)	99%
Wheel Speed (kph)	11.45%	0.47	<=55.61 (3/4)	47%
Front Brake Activation (0/100)	0.60%	0.03	100	71%
	C_2 (36.98%)			
Steering (%)	73.21%	1.00	<=-0.94 (2/4)	52%
Wheel Speed (kph)	54.71%	0.75	<=19.12 (1/4)	53%
Front Brake Activation (0/100)	19.69%	0.27	100	95%
Throttle (%)	13.24%	0.18	<=5.56 (1/2)	100%
Rear Brake Activation (0/100)	0.94%	0.01	100	77%
	C_3 (19.37%)			
Throttle (%)	78.40%	1.00	>5.56 (2/2)	91%
Rear Brake Activation (0/100)	61.63%	0.79	0	100%
Front Brake Activation (0/100)	51.62%	0.66	0	99%
Wheel Speed (kph)	37.13%	0.47	<=55.61 (3/4)	91%
Steering (%)	22.76%	0.29	<=1.84 (3/4)	98%

[1] *Values in the parentheses are the ranges of variable discretization the modal value belongs to.*

In the S_2 group (22% of sample), the rank order is different; front brake activation is the prevailing variable, in terms of influence to the knowledge of the dependent variable's state (S_2), followed by wheel speed, throttle and rear brake activation; in the second group of actions at the beginning of the incident, steering is the least influential variable. Actions belonging to the S_2 group show that the rider most likely will use the front brake in low speed, again by executing a minor maneuver. Finally, there is a third group S_3 (19% of cases) where throttle ranks first in terms of influence to the target variable S_3, followed by rear brake activation, steering and wheel speed; the front brake activation is not significant. Actions belonging to the S_3 group characterize accelerations in high speed where the rider executes a minor maneuver; brakes are not activated. A more thorough look at the results show that, although S_1 and S_2 cannot be easily intuitively assigned to specific riding situations, S_3 seems to be a characteristic action met in overtaking, as the rider suddenly accelerates in high speed and executes a minor maneuver.

Similarly, rider's actions during an incident can be clustered into three groups (Table 3). The first group C_1 encompassing the highest percentage of cases (44%) is more influenced by steering, rear brake activation, throttle and wheel speed rather than the front brake activation. This means that C_1 encompasses actions during which the rider executes a minor maneuver, while brakes are activated in medium speeds. In the second group C_2 (37% of cases), as in the case of C_1, steering is again critical, but wheel speed is more influential than braking. Moreover, front brake rather than rear brake is more influential. C_2 describes actions there the rider executes a minor maneuver in very low speed when both brakes are activated. Finally, C_3 is strongly related to throttle and the non-activation of both brakes, whereas wheel speed and steering is less influential. This means C_3 involves actions where the rider accelerates in medium speed.

Interestingly, although S_i set of actions uniquely appear in each incident (each incident have one appearing S_i action), an incident will most likely involve more than two actions C_i actions during its evolution. The manner S_1 and C_i are related, as well as the manner a sequence of actions C_i during an incident may be formed with respect to the type of incident will be further investigated in the next section.

PREVAILING RIDING ACTIONS OF RIDING INCIDENTS

Bayesian classifier is developed in order to relate each riding action at the beginning of an incident to the specific incident categories, as well as the actions to follow during the incident. Table 4 shows the data specifications with respect to the dependent variable and independent variables. As can be observed, the dependent variable is an incident categorization and is general enough to encompass a significant amount of situations met by PTW riders on roads. The dependent variable results from observing all detected incidents through video recordings. As independent variables, apart from the action taken at the beginning of the incident (*Start*), four other variables are considered. The first three are binary variables $\{0,1\}$ and refer to whether the action C_1, C_2 or C_3 has been observed during the incident. The fourth variable quantifies the number of actions a rider may do during a specific incident.

TABLE 4: Data specification for the BN classifier.

Independent	Dependent
Start (S_1, S_2, S_3)	Incident Category
Nr of actions (*1*, *2*,…)	• Stationary obstacle,
Action C_1 (0 if no, 1 if yes)	• Moving obstacle,
Action C_2 (0 if no, 1 if yes)	• Overtake,
Action C_3 (0 if no, 1 if yes)	• Opposing Traffic.

The Bayesian classifier developed can use the associated data (steering, throttle, front/rear brake activation and speed) for the prediction incident type with relative high precision (function of the number of correct predictions of the target variable) of 82%. Results on the discovered associations are seen in Table 5 with respect to the binary relative significance ($\frac{\text{mutual information}_i}{\max\{\text{mutual information}_i\}}$, $i=1,2,...,n$ where n is the number of variables describing traffic conditions) of each input variable to the knowledge of the transitions and the modal value. Table 5 summarizes the classification results.

As seen in Table 5, in incidents involving moving obstacles or opposing traffic the rider, at the beginning of the incident, will most likely activate the rear brake in medium speed, executing a minor maneuver (S_1). The riders will then conduct on average 2 actions to avoid a crash, with the prevailing action being C_2, which is to activate both brakes and conduct a minor maneuver in very low speed.

TABLE 5: Results from associating the type of the incident with specific riding actions at the occurrence and during the incident.

Node	Binary relative significance	Modal Value		Node	Binary relative significance	Modal Value	
Moving Obstacle (28.57%)				Opposing Traffic (28.57%)			
Start	1	S_1	90%	Start	1	S_1	90%
Actions	0.542	2	40%	C_1	0.843	0	100%
C_3	0.067	0	90%	Actions	0.768	1	70%
C_2	0.014	1	70%	C_2	0.180	1	80%
C_1	0.008	0	80%	C_3	0.074	0	90%
Stationary Obstacle (28.57%)				Overtake (14.29%)			
Start	1	S_2	60%	Start	1	S_3	80%
Actions	0.518	2	70%	C_3	0.591	1	80%
C_1	0.336	0	50%	Actions	0.326	3	60%
C_3	0.280	0	100%	C_2	0.078	0	60%
C_2	0.012	1	60%	C_1	0.001	0	80%

In the case of incidents involving stationary objects, the most probable initial action is the front brake activation in low speed (S_2); the rider will most probably activate both brakes and conduct a minor maneuver in very low speed (C_2) during the course of the incident. Finally, overtakes are strongly related to the initial action S_3, that is accelerating at high speed, while conducting a minor maneuver. During the incident, there is a high probability of accelerating and executing 3 actions with the prevailing action being the acceleration (C_3).

CONCLUSIONS

Risk hindering in the behavior of riders is a common consideration in PTW safety. Until now, riders' behavior has been systematically studied through survey questionnaires and police reports, methods that may be biased, lacking critical information or encompassing errors and inaccuracies due to perception. Although current technological advances have fostered the conducting of naturalistic experiment that may provide very detailed information on the manner a rider behave on the road, little is still known on the manner a rider reacts to the emergence of a critical incident. The present paper proposes a methodology based on Bayesian Networks for identifying the riding behaviors that arise at the emergence of a critical incident based on high resolution monitored riding data (100Hz) consisting of information on wheel speed, throttle, steering, brake activation and associate them with typical types of incident such as incident involving a moving or stationary obstacle, overtake and incident involving traffic in the opposite direction of travel.

Based on previous results which have shown that deviations from the mean riding behavior may efficiently be related to incidents with different levels of criticality with respect to rider's risk *(18)*, different behavioral patterns describing actions that govern the manner a rider reacts to external stimuli are revealed both for the onset of incident and during its duration. These patterns are mainly described by the interrelations between the riding variables that are related to the mechanical characteristics of the PTW, such as front and rear braking activation, throttle position, steering angle and wheel speed. Furthermore, the proposed methodology efficiently relates the observed patterns with four rough riding situations/incidents that are characterized by different initial actions and by different likelihood of actions undertaken by the rider during the incident. These four riding patterns relate to the occurrence of moving or stationary obstacle, overtaking and the opposing traffic.

The proposed methodology is purely probabilistic and compatible with the uncertainty hindering in the rider's behavior. The revealed riding patterns may be explicitly distinguished. The latter associated with the fact that a very broad categorization of observed incidents has been proposed, results to a flexible characterization of riding behaviors at the emergence and during and incident. Further research is needed on the variability of the observed behavioral patterns across different riders and different riding settings.

ACKNOWLEDGMENTS

The data used in this paper is part of the work conducted during the focused research collaborative project "2-BE-SAFE – 2-Wheeler Behavior and Safety" co-funded by the European Commission (www.2besafe.eu).

REFERENCES

1. Mannering F.L. and Grodsky L.L., Statistical analysis of motorcyclists' perceived accident risk, Accident Anal. Prevent. 27 (1995), pp. 21–31.
2. Jonah BA, Thiessen R, Au-Yeung E, 2001. Sensation seeking, risky driving and behavioral adaptation. Accid Anal Prev, 33:679–684.
3. Horswill M. S. and Helman S. (2003). A behavioral comparison between motorcyclists and a matched group of non-motorcycling car drivers: factors influencing accident risk, Accident Analysis and Prevention 35, 589–597
4. Harrison W.A. and Christie, R 2005, Exposure survey of motorcyclists in new South Wales, Accid. Anal. Prev. 37 (2005), pp. 441–451
5. Chen, C.F., (2009). Personality, safety attitudes and risky driving behaviors—evidence from young Taiwanese motorcyclists. Accident Analysis & Prevention 41, 963–968.
6. Chen C.F. and Chen, C.-W. (2011). Speeding for fun? Exploring the speeding behavior of riders of heavy motorcycles using the theory of planned behavior and psychological flow theory, Accident Analysis & Prevention, 43(3), 983-990.
7. V. Cossalter and R. Lot, "A motorcycle multi-body model for real time simulations based on the natural coordinates approach," Vehicle System Dynamics, vol. 37, no. 6, pp. 423–447, 2002.
8. L. Nehaoua, S. Hima, N. S´eguy, H. Arioui, and S. Espi´e, "Design of a new motorcycle riding simulator: Open-loop tests," American Control Conference, New York City, USA, july 11-13, 2007.
9. Liu, C. C., Hosking, S. G., & Lenné, M. G. (2009). Hazard perception abilities of experienced and novice motorcyclists: An interactive simulator experiment. Transportation Research Part F, 12(4), 325-334.
10. S. Hima, L. Nehaoua, N. Séguy and H. Arioui, Motorcycle Dynamic Model Synthesis for Two Wheeled Driving Simulator. To appear in International IEEE conference on intelligent transportation systems, 2007.
11. Jovanis P. P., Valverde J. A., Wu K.-F. Shankar, V. (2011). Naturalistic Driving Event Data Analysis: Omitted Variable Bias and Multilevel Modeling Approaches, TRB 90th Annual Meeting Compendium of Papers DVD.
12. NHTSA (2006), The 100-Car Naturalistic Driving Study Phase II – Results of the 100-Car Field Experiment, www.nhtsa.dot.gov, USA.

13. Csepinszky, A. and Benmimoun, M. (2010). An operational perspective on the organisation of large scale field operational tests of intelligent vehicles, András Csepinszky, ERTICO - ITS Europe, ITS World Congress, Busan, 25-29 October.
14. SHRP2 (2011). A Foundation for Safer Driving, April, Transportation Research Board, Washington DC US.
15. Regan, M., Triggs, T., Young, K., Tomasevic, N., Mitsopoulos, E., Stephan, K. and Tingvall, C. (2006) On-road evaluation of Intelligent Speed Adaptation, Following Distance Warning and Seatbelt Reminder Systems: final results of the TAC SafeCar project, Monash University Accident Research Centre, Report No. 253.
16. FESTA (2008), Field opErationalteStsupporT Action (FESTA) - D2.4 Data analysis and modelling, FESTA project, European Commission, Belgium.
17. 2BESAFE (2009). Literature review of data analysis for naturalistic driving study, deliverable 4, December, 2BESAFE project, European Commission, Belgium.
18. Vlahogianni, E. I., Yannis, G., Golias, J. C., Eliou, N. and Lemonakis, P. (2011). Identifying Riding Profiles Parameters from High Resolution Naturalistic Riding Data, Proceedings of the 3rd International Conference on Road Safety and Simulation (RSS2011), September 14-16, Indianapolis, Indiana, US.
19. Cheng J. and R. Greiner. Learning Bayesian Belief Network Classifiers: Algorithms and System, AI 2001, LNAI 2056, 2001, pp. 141-151.
20. Friedman, N., Geiger, D. & Goldszmidt, M., (1997). Bayesian Network Classifiers. Machine Learning, Vol. 29, 131–163. Kluwer, Boston.
21. Vlahogianni, E. I., Webber, Ch. L. Jr., Geroliminis, N. and Skabardonis, A. (2007). Statistical Characteristics of Transitional Queue Conditions in Signalized Arterials, Transportation Research Part C, 15(6), 345-404.
22. Castillo E., Menendez J. M. and Sanchez-Cambronero S. (2008). Predicting traffic flow using Bayesian networks, Transportation Research Part B: Methodological, 42(5), 482-509.
23. Vlahogianni E.I., Karlaftis, M.G., Golias, J. C. and Halkias, B. (2010) Freeway Operations, Spatio-temporal Incident Characteristics, and Secondary-Crash Occurrence, Transportation Research Record: Journal of the Transportation Research Board, 2778, 1-9.
24. Dempster A. P., Laird N. M., and Rubin D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, B39:1–38.
25. Barnett, V. & Lewis, T. (1994). Outliers in Statistical Data, 3rd edn. John Wiley & Sons.
26. 2BESAFE (2010). Design of a naturalistic riding study-Implementation plan, Deliverable 5 of work package 2.2, 2BESAFE project, European Commission, Belgium.