U&I Aware: A Framework Using Data Mining and Collision Detection to Increase Awareness for Intersection Users

Flora Dilys Salim[i], Seng Wai Loke[ii], Andry Rakotonirainy[iii], Shonali Krishnaswamy[i]
[i]Caulfield School of Information Technology, Monash University, Australia.
{Flora.Salim, Shonali.Krishnaswamy }@infotech.monash.edu.au
[ii]Department of Computer Science and Computer Engineering, La Trobe University, Australia.
s.loke@latrobe.edu.au
[iii]Centre for Accident Research and Road Safety Queensland, Australia. r.andry@qut.edu.au

Abstract

An intersection safety system should adapt to the particular characteristics that identify an intersection, by mining traffic and collision data. Given the large amount of sensor data that are obtained for intersections and from sensor-equipped cars, analysis and learning of such data is essential. This paper presents a new method to improve safety at intersections using a combination of a mathematical based collision detection algorithm and data mining. A number of scenarios at a simulated intersection are explored with encouraging results from our data mining implementation. The results suggest that our approach can help improve situation awareness and automate understanding of intersections, which, in turn, can be used to increase safety at intersections.

1. Intersections at Stake

The annual toll of human loss caused by intersection crashes has not significantly changed in more than 25 years, regardless of improved intersection design, more sophisticated cars, and more sophisticated ITS technology [1]. The high accident and fatality rate at intersections is due mainly to the complexity of each intersection. Therefore, the complex nature of intersection collisions requires systems that warn drivers about possible collisions. In addition, given the uniqueness of each intersection, rather than manually fine-tuning a system for each intersection, an intelligent system for intersection safety should be able to adapt to different types of intersections automatically [2].

This paper aims to investigate the issues in constructing intersection collision warning systems and discuss how the algorithm to detect potential collisions can be improved by means of preselection and data mining. We then validate our findings through an implemented simulation. First, Section 2 discusses related works in intersection collision warning and/or avoidance systems and robotic collision avoidance. Section 3 discusses how data mining, including ubiquitous data mining, can be useful for extracting knowledge concerning intersections. Section 4 discusses implementation issues, and solutions for the effectiveness and efficiency of collision detection in road intersections. Section 5 concludes the paper.

2. Related Work

2.1. Intersection Collision Warning and/or Avoidance Systems

There have been a number of initiatives in developing intersection collision warning systems and/or avoidance systems. Currently, no existing intersection collision warning and avoidance systems can tackle intersection collision problems entirely. Intersection collision warning and avoidance systems can be categorized as either *vehicle-based*, *infrastructure-only* or as *infrastructure-vehicle cooperative* [2]. Many existing Intersection Collision Warning Systems are still infrastructure-only systems, and are limited in certain aspects, as described in [2]. Vehicle-based intersection collision warning systems are fairly effective for a single vehicle [2]. However, in an intersection, the potential danger normally impacts more than one vehicle, therefore, a cooperative system is preferred [2]. However, to our knowledge, existing research projects in cooperative systems for intersection safety do not mention techniques to discover crash patterns and pre-crash behaviour

associations [2], which are essential to detecting and reacting to potential threats. A generic framework that can automatically adapt to different intersections is required for efficient deployment; however, these projects have not addressed this issue [2].

2.2. Robotic Collision Avoidance

Studies in robotic collision avoidance have existed for many years [3], [4], [5]. Robots need to be able to find their own way to their destination as well as to avoid obstacles on their path. Although it seems that robotic collision avoidance has much resemblance to the problem of road collision avoidance, those two subjects differ in many aspects, which are as follows:

1. Robotic collision avoidance mostly focuses on static obstacles, such as walls. There are only a few researches in the area that consider dynamic obstacles. Whereas in road collision avoidance, we deal mostly with dynamic obstacles; therefore, the movement attributes of all objects must be taken into account, and a knowledge base should be updateable.
2. Robotic collision avoidance focuses on the goal of the robotic tasks such as to find a way out of a room. Road intersection collision avoidance focuses on getting to the destination safely. Therefore, having a direction might not be necessary for robots, as long as their goals are achieved.
3. Robotic collision avoidance is a component in the path finding [3], [4]. In another words, a path is the outcome of a collision avoidance process. However, in road collision avoidance, intended path of the driver is known to a certain extent by using sensors and is decided before a collision avoidance process.
4. Road collision avoidance does not require full automation such as robot collision avoidance. The output of road collision avoidance is primarily warning to drivers or other road users, whereas robot collision avoidance requires autonomous actions.

Due to the above differences, we need to approach road collision avoidance issues differently from robotic collision avoidance. However, there is a dynamic knowledge base technique introduced by [5] that can be adapted to road collision avoidance. A static knowledge base contains static rules of obstacles detection and collision avoidance. No new rules added to the knowledge base. Dynamic knowledge base involves learning to accumulate and refine rules in knowledge base to adapt to situational changes. As situations in road intersections, such as traffic trends, weather changes, and collision patterns, are very dynamic and vary from one intersection to another, a dynamic knowledge base should be employed.

Pervasive computing technology and intelligent systems provide significant potential for improving intersection safety. Our previous paper [2] noted that the safety of intersections can be improved by integrating the advances in sensors development, Intelligent Transportation Systems and situation-based reasoning. Intelligent agents and context-awareness have been integrated into ITS, such as for smart autonomous cars and traffic monitoring.

3. The need for data analysis and learning

Since there is considerable amount of data from the in-vehicles and roadside sensors, it is essential to make sense of the sensor data via data analysis techniques. There have been a number of research projects on data mining in the area of ITS, such as for driver's behavior recognition, traffic optimization, and incident detection [2]. The Pantheon Gateway Project [6] detects real time changes in traffic conditions (speed, volume, occupancy) using real time highway data. Traffic condition changes, present accidents, and special events that affect the traffic can be detected in real-time based on the learnt traffic patterns. Data mining is proven to be effective for extracting traffic patterns and trends.

Currently, collision warning systems mostly react to events that might cause collision. Intersection collision warning systems should also evolve by adapting to information gained from analysis of sensor and historical data in the intersection. By learning from historical data of collision and near-collision events, improved detection and reactive behaviour can be achieved since the knowledge base of the intersection is evolving in the U & I Aware (Figure 1). With in-vehicle learning (i.e. classification and association) of driver behaviours, the intersection collision warning system can be informed when a driver exhibits dangerous driving behaviours. When the result of learning of historical collision, traffic, and driver behaviour data in the intersection is integrated into the system, the system can gain better knowledge of the intersection for better crash prediction. The following section discusses cases in intersection where mere mathematical based algorithms is insufficient and learning of sensor data in intersection is useful to improve knowledge of the intersection and to recognize dangerous situations and hazards at the intersection.

4. Capturing useful knowledge with data mining: going beyond simple trajectory analysis

We have implemented an algorithm that can be used to detect collision [7] by calculating a future collision point (x_+, y_+):

$$x_+ = \frac{(y_2 - y_1) - (x_2 \tan\theta_2 - x_1 \tan\theta_1)}{\tan\theta_1 - \tan\theta_2} \; ; \;$$

$$y_+ = \frac{(x_2 - x_1) - (y_2 \cot\theta_2 - y_1 \cot\theta_1)}{\cot\theta_1 - \cot\theta_2}$$

, where θ is the angle between the horizontal line and car trajectory.

The time for each car to reach the future collision point (TTX) [7] is calculated by:

$$TTX_1 = \frac{|\vec{r}_+ - \vec{r}_1|}{|\vec{v}_1|} sign((\vec{r}_+ - \vec{r}_1).\vec{v}_1)$$

$$TTX_2 = \frac{|\vec{r}_+ - \vec{r}_2|}{|\vec{v}_2|} sign((\vec{r}_+ - \vec{r}_2).\vec{v}_2)$$

where v is velocity of each car and r is the vector of the coordinate (x, y) [7].

The algorithm, which is based on a vector formula and implemented in a computer based simulation of intersections, is able to predict a near exact time and point of collision. The simulation parameters have been instrumented to mirror real world situations. Based on the above formula, the information required to calculate a collision point and the time to collision are speed, vehicle size, travel direction, current position, and angle. Table 1 lists possible sensors that can be used to acquire the required information.

**Table 1. List of Sensors Used to Capture Traffic Data**

Required Information	In-vehicle sensors	Roadside sensors
Speed	Speedometer	Camera, Inductive loop detector, Traffic-Dot [8]
Vehicle Size	Built-in information	Traffic-Dot [8]
Travel Direction	Camera	Camera
Current Position	GPS, GIS	Camera, Inductive loop detector, Traffic-Dot [8]
Angle	Camera, Steering Wheel	Camera, Inductive loop detector, Traffic-Dot [8]

We implemented the Ubiquitous Intersection Awareness (U & I Aware) framework, which aims to achieve holistic situation recognition at road intersections. We use a computer based simulation of two different scenarios: intersection with traffic lights and without traffic lights. At this stage, computer based simulation is an acceptable proof of concept, since the scenarios that we implement involve collisions that are difficult to be simulated in the real world due to the constraint of resources and technology. The simulation parameters are as follows:

1. Intersection module: intersection type, leg, lane, lane group, traffic control.
2. vehicle: speed, acceleration, size, type, position, angle, maneuver.
3. driver: profile, intended destination, choices of maneuver.

The vehicles are randomly generated with different speeds, maneuvers, position and trajectory at the end of each intersection leg. Each vehicle is should observe the traffic light signals, safe following distance (3 seconds), safe stopping distance (2 seconds), and the speed limit. Random "naughty" vehicles are generated in the simulation to test the ability of the collision detection and learning algorithms.
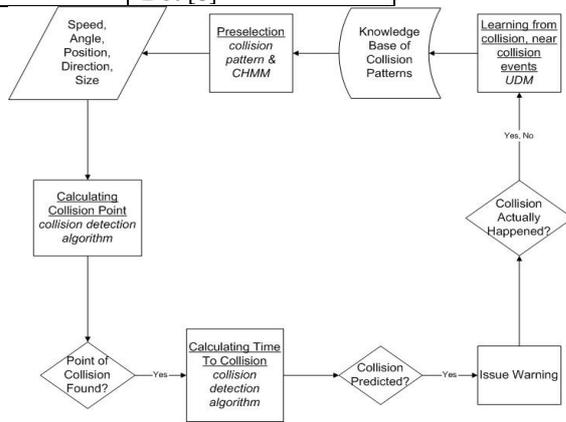
[fit this into one column]

Figure 1. U & I Aware Framework Strategy

Based on the simulation, while the technique can be used for collision warning, we found that there are two limitations of such an approach:

1. It requires calculation for each possible pair of vehicles in the intersection; therefore, computational cost is high. Collision detection must be performed each time a car moves from its current position. Each time a vehicle moves, it must know about possible collisions with any other vehicles in the vicinity.
2. It is unable to capture new and useful knowledge about collisions or near collision events that can be used to enhance the prediction of collision.

In the following two subsections, we will discuss these two limitations and how each can be addressed.

4.1. Improving the efficiency of collision detection

In dealing with the first issue mentioned, we implemented a preselection algorithm, so that collision detection is only performed on pairs of cars that have the possibility of collisions based on the known intersection collision patterns. Preselection is implemented by choosing only the vehicles that exhibits behaviours, location, and driving manoeuvres that match the collision patterns in the knowledge base. The crash pattern knowledge base is implemented as a hash table filled with crash pattern class objects. Each crash pattern consists of a name, a manoeuvre, a direction, an intersection leg location, and a delegate function to find conflicting direction and manoeuvres. For example, given there is a cross intersection and the knowledge base contains a collision pattern named "perpendicular paths", which means collision normally happens between vehicles that have straight manoeuvre movement when entering the intersection, their conflicting paths will intersect at an angle of around 90 degrees. If a car enters the intersection from the south leg of the cross intersection detection with a straight manoeuvre movement, collision detection will be performed on this car against every other car that is currently located on perpendicular paths (i.e., west and east legs of the intersection), or moving straight towards the intersection. Therefore, performance is improved by not needing to check every pair of cars at the intersection for possible collision. Predicting

intended driving manoeuvres can be done one second before the actual manoeuvre takes place [9]. We enumerate the results of the manoeuvre prediction [9], which can predict the following manoeuvres: passing, turning right, turning left, changing lane right, changing lane left, starting, and stopping, in our knowledge base of collision patterns. Each collision pattern consists of one of those manoeuvres that can possibly predicted by in-vehicle sensor implementation and CHMM (Figure 1). The efficiency of the collision detection algorithms [7] has been improved using our preselection method, as seen in Figure 2.

The implementation of the preselection algorithm is described as follows:

1. The knowledge base of the intersection in this example records two types of side collision patterns: perpendicular left with straight manoeuvre and perpendicular right with straight manoeuvre.
2. Each car that moves needs to be checked for side collision prediction. However, we will not compare each car to every other car in the intersection. Only cars that are located within a certain area and exhibiting certain manoeuvres are selected. As for the truck B located at the right leg of the intersection in Figure 2, the algorithms will only be applied on vehicles on the upper and bottom legs that are exhibiting straight manoeuvre, based on perpendicular left with straight manoeuvre and perpendicular right with straight manoeuvre patterns. Those vehicles are car A at the bottom leg and car C at the upper leg.
3. Only after preselection is executed, only then the pair-wise collision detection algorithm is applied.

4.2. Improving the effectiveness of collision detection

To tackle the inability of mathematical algorithm to capture new knowledge, learning is performed by using classification and association rules of data mining. New events are matched with the existing classes in the patterns repository of the intersection central agent or the car agent, depending on where learning happens. If a collision happens outside a known pattern, a learning process can detect and add a new collision pattern.
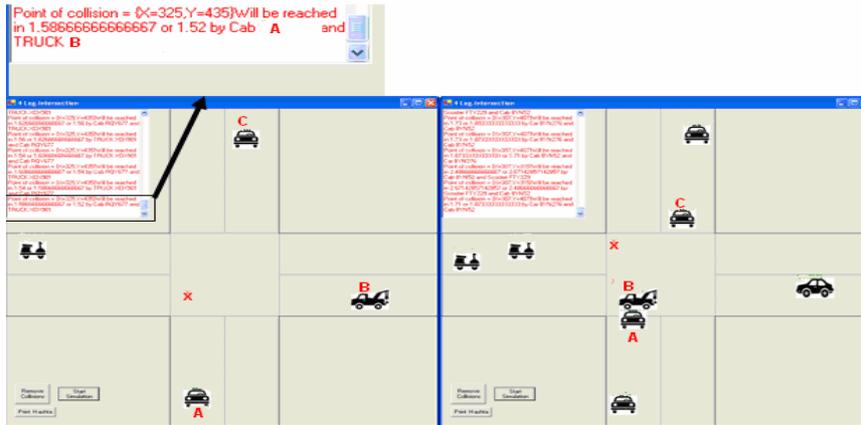
Figure 2. Cross Intersection without Traffic Lights Implementation

There are a number of improvements and enhancements that can be added to the plain collision warning system that is based only on trajectory calculations. These are done via mining of data assumed to be obtained from on the road sensors and in vehicle sensors in order to characterize: collision patterns, normal conditions of intersection within collision-free periods, normal behaviours of drivers within collision-free periods, abnormality in intersection that leads to collision (as antithetical models of the models of normal behaviours above), and red light running:

Collision patterns. Collision patterns in an intersection can be learnt when there is data about vehicle manoeuvre, direction, and angle. We have assumed these data in our simulation. Near-collision event is determined by a parameter of distance between two vehicles that almost collide to each other. Whenever there is a collision or near-collision event in our intersection simulation, data from colliding pair of vehicles are collected and mined. The data has six attributes, which consists of three attributes (i.e. direction, manoeuvre, and angle) from each colliding vehicle pair. We have implemented this with C4.5 decision tree (J48 classifier [10]) and the vehicle direction attribute is nominated as the class. It successfully classifies perpendicular crashes in cross intersection. The results also exhibit the most common crash patterns in a particular intersection where traffic data are learnt. For example, using randomly seeded data, our results show that in this particular intersection, where data were gathered and learnt, vehicles that travel with straight manoeuvre from the left leg to the right leg of intersection tend to collide with vehicles that travel with straight manoeuvre from the lower leg to the upper leg (Figure 3).
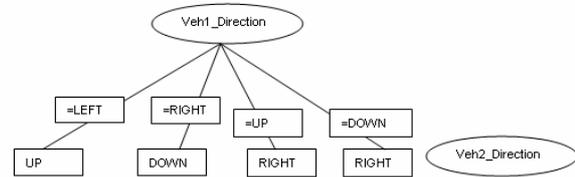

Figure 3. Collision Patterns based on Vehicle Direction classified with C4.5

To realise all the possible crash patterns in an intersection, Bayesian Network classifier [10] is used to classify the same data. The classification shows all the possible collision patterns that might happen with the probability rate of each crash pattern (Figure 4).

Veh2_Direction	UP	DOWN	RIGHT	LEFT
UP	0.014	0.014	0.042	0.931
DOWN	0.022	0.065	0.891	0.022
LEFT	0.583	0.25	0.083	0.083
RIGHT	0.611	0.278	0.056	0.056

Figure 4. Collision Patterns based on Vehicle Direction classified with Bayesian Network (simulated with Weka [10])

The crash patterns enumerate four possible original leg position and direction. Vehicles that travel from the right leg to the left leg have the highest probability of crash with vehicles from the lower leg to the upper leg. This result conforms to the result of classification with C4.5. Note that these results were obtained from our simulated data for one intersection. Applying the same technique to a different intersection (with different data) could lead to different likely situations for collisions – the point is that applying such learning techniques would enabling such collision situations to be recognized automatically and identified as "dangerous" patterns.

Normal conditions of intersection within collision-free periods. Other situations that can be learnt with data mining are traffic and drivers' attributes within

collision-free periods. Attributes of intersection's traffic that can be monitored for a period of time are as follows: period, number of collisions, traffic volume, time of day, peak hours (yes / no), average speed of vehicles, safe ranges of vehicle speeds. These attributes are learnt with C4.5 classification to characterise incident free behaviours at intersections, safe thresholds (i.e. the attributes' values under safe or normal situations), and then used to identify hazardous situations at the intersections when possibilities of collisions are present.

Normal behaviours of drivers within collision-free periods. Attributes of drivers that are used to characterise an ideal driver, who is risk-free at the intersection, are as follows: above speed limit, above average speed, collision presence, approaching intersection, distance to intersection, increasing speed (yes / no). Once we know the acceptable threshold of those attributes for a collision free drive in an intersection, we would be able to easily identify abnormality in intersection or abnormal behaviours of drivers that are using the intersection if they exhibit any attribute that exceeds the acceptable threshold. We implemented a small random number of "naughty cars" in our crossroad simulation, where some of them would violate the red light and the others would have speed below the limit. One result of learning driver behaviour data using C4.5 is that most cases of speed limit violation occur if the vehicle increases its speed when leaving the approach leg (any intersection leg where there is incoming traffic to the intersection centre) and entering the centre of the intersection.

Red light running. Red light running behaviour is not a pattern, but an event that can happen in any pattern of collision. Therefore, it is essential to keep monitoring such events, which can possibly occur when:

1. a vehicle speeds up on yellow signal and hasn't been able to reach the opposite leg when the traffic signal turns to red,
2. a vehicle tries to stop on yellow signal, but not slow enough to stop behind the current leg's stop line, and
3. a vehicle starts early when the red light is about to turn green.

Therefore, the attributes that are learnt to predict red light running behaviours are: change of speed, distance to stopping line, current traffic light colour, and traffic light change period.

6. Conclusion

On top of our implementation of collision detection based on computing trajectories, we desire that the whole system at each intersection evolve over time through better situation awareness and understanding for better safety. The contribution of this paper is a novel hybrid model to improve safety at intersections using a combination of a mathematical based collision detection algorithm and data mining, which we have demonstrated, via simulation, to be effective and feasible. Simulation is used as our work involves what might be termed "futuristic" yet currently realizable scenarios that we envision will be deployed in the future, if not increasingly so. This paper proposes a framework that existing technologies (i.e., hardware, sensors, etc), as we pointed out, can support, and we aimed to first demonstrate by simulation our approach before actual deployments can be considered.

7. References

[1] U.S. Department of Transportation – Federal Highway Administration, Institute of Transportation Engineers, *Intersection Safety Briefing Sheet*, April 2004; http://www.ite.org/library/IntersectionSafety/BreifingSheets.pdf.

[2] F. D. Salim, S. Krishnaswamy, S. W. Loke, A. Rakotonirainy, Context-Aware Ubiquitous Data Mining Based Agent Model for Intersection Safety, *Proc. of EUC 2005 Workshops*, 2005.

[3] C. Fayad and P. Webb, "Optimized Fuzzy Logic Based on Algorithm for a Mobile Robot Collision Avoidance in an Unknown Environment", *Proc. Of 7th European Congress on Intelligent Techniques & Soft Computing*, Aachen, Germany, September 13-16, 1999.

[4] D. Fox, W. Burgard, and S. Thrun, "The Dynamic Window Approach to Collision Avoidance", *IEEE Robotics & Automation Magazine*, vol. 4, no. 1, 1997.

[5] S. Manivannan, "Robotic collision avoidance in a flexible assembly cell using a dynamic knowledge base", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 23, issue 3, May-June 1993, pp. 766 - 782

[6] R. L. Grossman, M. Sabala, J. Alimohideen, A. Aanand, J. Chaves, J. Dillenburg, S. Eick, J. Leigh, P. Nelson, M. Papka, D. Rorem, R. Stevens, S. Vejcik, L. Wilkinson, and P. Zhang, "Real Time Change Detection and Alerts from Highway Traffic Data", *Proc. Of ACM/IEEE Supercomputing, 2005*. pp. 62 – 69.

[7] R. Miller and Q. Huang, "An Adaptive Peer-to-Peer Collision Warning System", *Proc. of Vehicular Technology Conference (VTC) Spring 2002,* Birmingham, Alabama.

[8] S. Coleri, S. Y. Cheung, P. Varaiya, *Sensor Networks for Monitoring Traffic*, invited paper to Allerton Conference,

September 2004; http://www.eecs.berkeley.edu/~csinem/academic/publications/allerton_traffic_2005.pdf

[9] N. Oliver and A. Pentland, "Graphical Models for Driver Behavior Recognition in a SmartCar", *Proceedings of IEEE Intl. Conference on Intelligent Vehicles 2000*, Detroit, Michigan, October 2000.

[10] I.H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*, 2nd edition, Morgan Kaufmann, San Francisco, 2005.