SIGNS WITH SMART CONNECTIVITY FOR BETTER ROAD SAFETY

SOLUTION ARCHITECTURE

Abstract

The Safe System (SS) approach to road safety emphasizes safety-by-design through ensuring safe vehicles, road networks, and road users. With a strong motivation from the World Health Organization (WHO), this approach is increasingly adopted worldwide. Considerations in SS, however, are made for the medium-to-long term. Our interest in this work is to complement the approach with a short-to-medium term dynamic assessment of road safety. Toward this end, we introduce a novel, cost-effective Internet of Things (IoT) architecture that facilitates the realization of a robust and dynamic computational core in assessing the safety of a road network and its elements. In doing so, we introduce a new, meaningful, and scalable metric for assessing road safety. We also showcase the use of machine learning in the design of the metric computation core through a novel application of Hidden Markov Models (HMMs). Finally, the impact of the proposed architecture is demonstrated through an application to safety-based route planning.

Introduction

In its Global Status Report on Road Safety – 2015, the World Health Organization (WHO) noted that the worldwide total number of road traffic deaths has plateaued at 1.25 million per year, with tens of million either injured or disabled [1]. Different initiatives, such as the United Nations' initiative for the 2011-2020 Decade of Action for Road Safety, have led to improvements in road safety policies and enforcements. However, the WHO notes that the progress has been slow and has maintained the call for urgent action to reduce these figures [2].

Added to the losses in human lives and wellbeing, considerable monetary losses are incurred in medical expenses, infrastructure repair, and production downtime. While the worldwide figures have plateaued, the Global Status Report does indicate higher road fatalities and injuries in low-income countries. Such disparity, as noted in [3], signals a barring-limitation in low-income countries to improve road-safety by adopting solutions implemented in high-income countries.

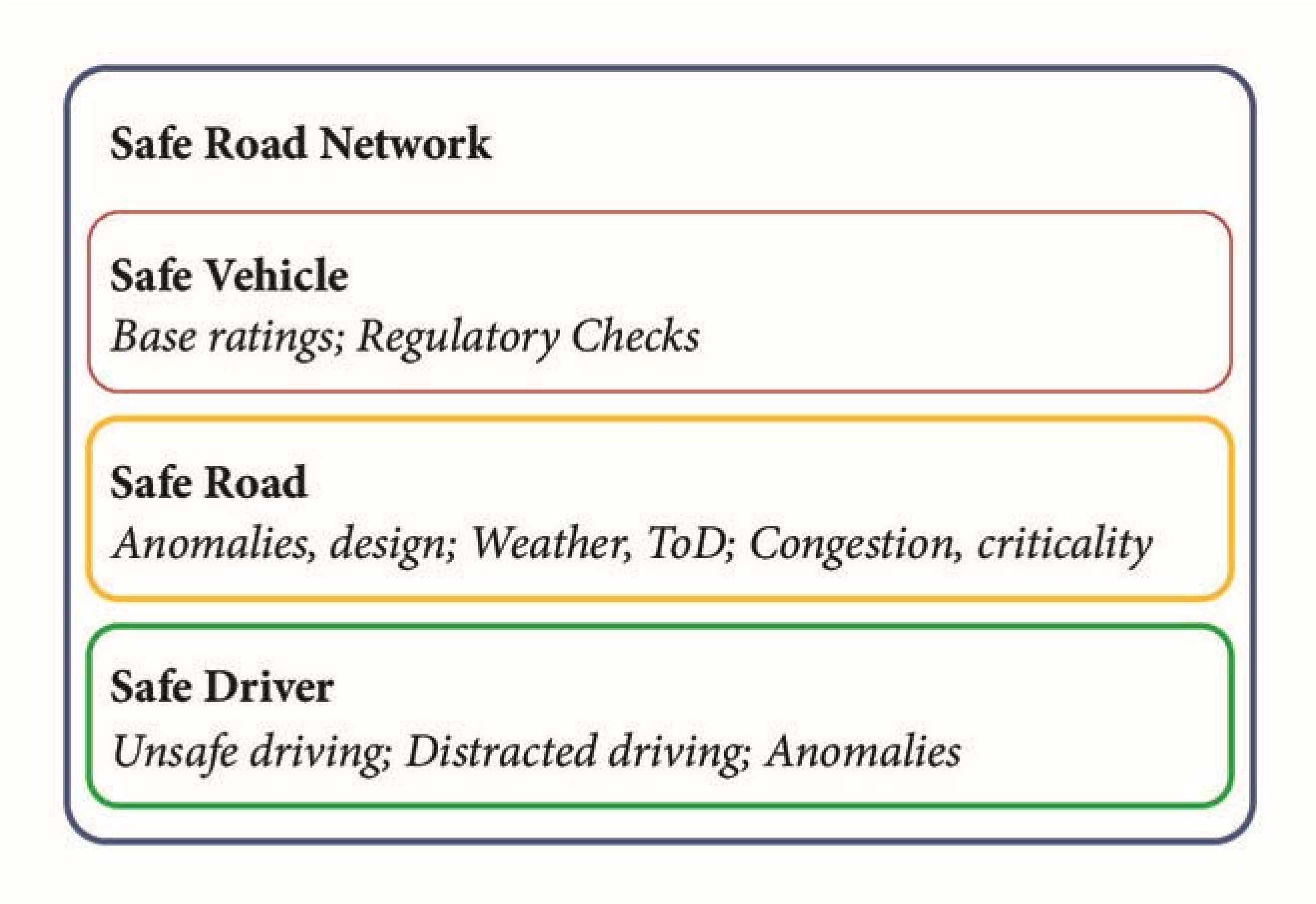
The WHO describes different measures that can be implemented with minimal economic impacts in its "Save LIVES: Road Safety Technical Package" [4]. A cornerstone of these steps is realizing economic systems for "monitoring road safety by strengthening data systems". Meanwhile, a key theme in the package is motivating the adoption of a Safe System approach, which is a holistic approach to road safety that parts from traditional management solutions by emphasizing safety-by-design.

An IoT Architecture for Assessing the Safety of a Dynamic Road Transport System

In reviewing the related works in the previous section, we showcased how various advances are enabling the assessment of safety of vehicles, roads, and drivers. The objective of this section is to introduce a novel and adaptive IoT architecture that enables the assessment of safety in a city's road network. We elaborate on the assessment elements and how they can be used to synthesize a single, meaningful indicator for safety. We also describe the architecture components and their interrelationships, including a robust computational core for safety assessment.

Assessment Elements

The way the SS approach comprises the three elements of safe vehicle, safe road, and safe driver facilitates a hierarchical safety assessment approach whereby the safety of the individual elements can provide a collective indicator of safety for the road network, as illustrated in Figure 3. In turn, this indicator can be concatenated from the assessment of individual road segments, to routes, to the road network.



Elements in assessing the safety of the road network based on the safe systems approach.

For vehicles, the assessment core would rely on inferences from the vehicle's Vehicle Identification Number (VIN) (thus establishing car make, manufacturing, and base safety rating); regulator's information (e.g., the outcome of the last regulatory check); and the updated information from an OBD-II unit.

For roads, the reliance on both historic and newly sensed data can facilitate an understanding of the road context, the road network map, and establish severity of turns, the presence of shoulders, and whether the road is on a hillside. Identifying a vehicle's location through an on-board GPS (whether dedicated or through a smartphone) would thereby provide for a first level of this understanding. Meanwhile, in-vehicle sensing can be utilized to report road anomalies (e.g., potholes). Sensing based on the OBD-II interface can also help identify instances of skidding, which in turn can be correlated with weather or reports. Weather, together with time of the day (ToD) and reports on the availability and/or state of road-lights, can be combined to ascertain visibility. Moreover, basic traffic information can indicate congestion levels, as well as road/segment criticality, i.e., repeated selection in route plans generated within a short span of time.

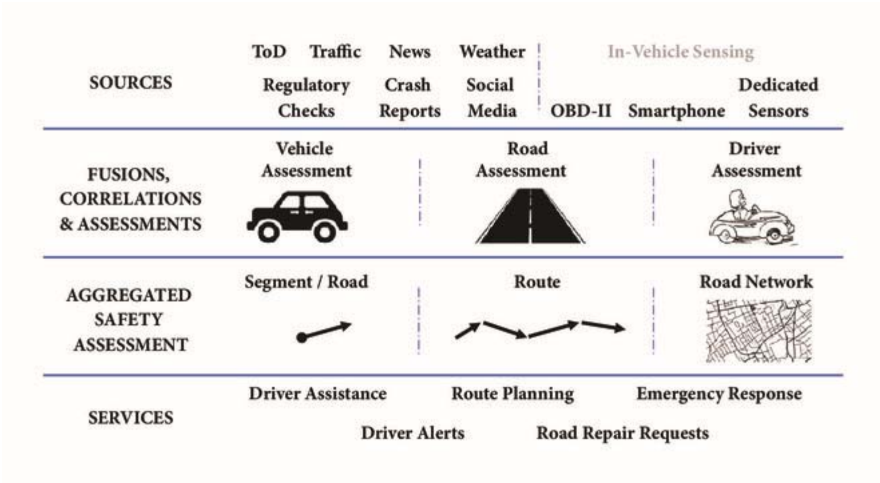
As for drivers, the growing maturity of DBM will facilitate identifying unsafe or distracted driving behavior. It will also enable the recognition of localized driving behavior anomalies, facilitating the identification of emerging incidents, e.g., several swift lane changes to avoid a new obstacle in the road due to a falling tree trunk or other road debris.

It is possible to consider a meaningful safety metric based on the live (or real-time) status of the road. For example, the safety level of a certain segment/road depends on the aggregate safety of vehicles currently traversing it, combined with the number of potholes and/or the wetness or how slippery is the road, in addition to safety/alertness of the drivers on the road.

In designing our architecture, we exploit three important dependencies. The first is between the SS elements, e.g., how well a car can handle a certain road, or how some drivers exhibit safer behavior in instances of higher visibility. The second dependency is in between consecutive segment/roads, especially in terms of traversing vehicles and drivers. The third dependency is like the second but is established in time. Abrupt changes in safety levels can thus be viewed as an anomaly (outlier) or inferred as indicator to a substantial change in the road context.

Architecture Considerations

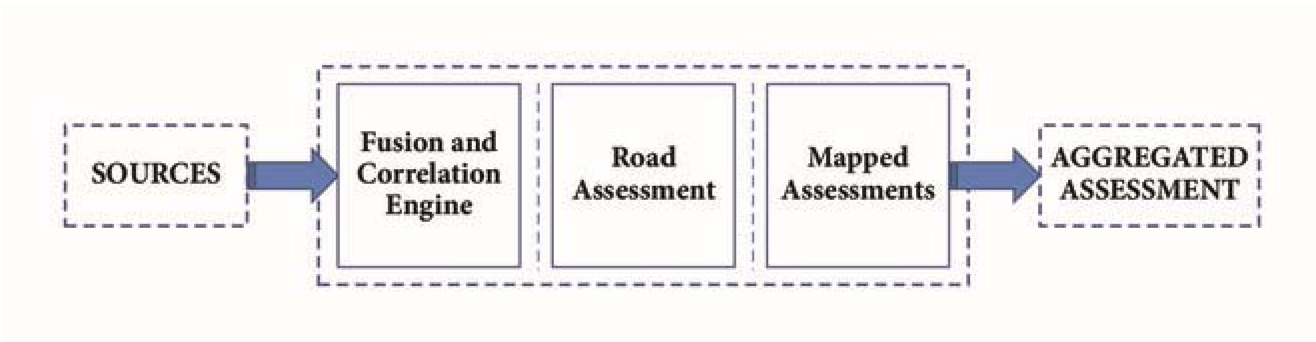
An illustration of architecture consideration is provided in Figure 4. At the core of the proposed architecture is a computational core that operates in two tiers. The first is concerned with dynamically assessing the safety of each aspect in SS. The second utilizes the outcome of these three assessments to generate a single rating for the road.



Architecture considerations for dynamic assessment of road network safety.

The individual assessments rely on a wide array of sources that generally fall into two categories. The first is sourced through in-vehicle sensing, while the second through sources that establish the general context of the road. These include traffic problems, regulatory checks, crash reports, processing of news and social media, weather updates, and time of day.

As aforementioned, inferences and assessments in each of the three elements of SS can utilize the same sources. This motivates the placement of the data fusion and correlation element in the architecture, as shown in Figure 5. In data fusion, sources that yield to the same knowledge synthesis are combined to enhance its reliability, while correlation is concerned with cotiming and colocation of sensed data, as well as cross-validation of input knowledge streams.



An example breakdown of fusion, correlation, and assessment for a Safe System component.

Based on the output of the three assessment modules, a live representation of a segment's safety to traverse can be made. We elaborate on a model for generating this representation below. Once the aggregated safety assessments have been generated, the values can be provided to several services that include driver-assistance modules and generating requests for road repair or emergency response.

In the following section we demonstrate how this assessment core can be utilized in a safety-based route planning application

In-Vehicle Sensing

A possible setup for in-vehicle sensing comprises three elements: an OBD-II dongle, a dedicated sensor, and a smartphone. We evaluated two OBD-II interface alternatives: munic.box's Munic3 4G-cat4-WBT dongle and ScanTool's OBDLink MX Bluetooth. Both performed as desired on the tested vehicles. The dedicated sensor was Aaronia's GPS logger, which combines sensors for GPS localizations, compass, accelerometer, gyroscope, and altimeter with pressure sensor. The logger does not have wireless communication capabilities but can be connected to a PC through USB. Meanwhile, recent iOS and android-based phones were utilized for both their communication and sensing caspabilities.

While an objective of the work is to realize a cost-effective solution for dynamic assessment, the purpose of the setup described above was designed for both exploratory and validation purposed. Where the sensing elements were able to connect to the cloud repository directly, e.g., as in the case of the munic.box dongle, or the smartphone, the connection was allowed. Otherwise, a connection was established indirectly through another device, e.g., the logger through the PC laptop or ScanTool's dongle through the smartphone.

For the backend, sensed data and be transferred and stored on a PostgreSQL 9.6 built on a Red Hat Enterprise Linux Developer OS. This setup would allow for an Apache web portal access, facilitating quick prototyping and ease of simultaneous data reports from multiple sources.

The Assessment Core

Road networks comprise a complex, living structure, and thus require care in modeling any of their aspects. This is especially the case in cities and high density urban areas where traffic dynamics becomes further involved and further engaged with other planes of city activity.

In remarking on dependencies above, we noted that a reasonable model to assess a road network's safety should mind the interelement dependencies in SS, as well as the dependencies in both space and time. It should also facilitate the quantification of these interdependencies without substantial overhead.

The design of such assessment core is elaborated upon in the next section.

Designing a Dynamic Assessment Core

Recent thrust into machine learning and its different variants (supervised, unsupervised, and reinforcement-based) has made addressing high processing demands in solving various problems possible. The operational requirements highlighted in Section 3.4 above position the target assessment core as a typical candidate for machine learning considerations. In what follows, we showcase an example modeling to the road network that can facilitate a robust and dynamic processing.

Hidden Markov Modelling (HMM) is a powerful statistical tool for modelling time-series systems that can be characterized to represent probability distributions over a sequence of observations. The tool thus lends itself easily to the nature of data gathering found in IoT and smart cities applications. It further stands as a potential base model for several machine learning approaches, including Bayesian or Mixture Density Network inferences.

Our interest herein is in a novel application HMMs to our problem. Specifically, a first-order, time-homogeneous, and discrete HMM is employed to identify the degree to which traversing a certain road link is safe, thereby realizing a safety metric.

The HMM at hand can be formally defined by the five-tuple , where is an array of links' states; is the emission symbols that characterizes observations per each state; is the states' transition probabilities; δ is the emission (or output) probability matrix; and π is the initial state distribution array. We assume that the road link state in terms of safety, denoted , is hidden from the observer. We also assume that the current hidden state of a certain road link depends only on the preceding state of the same road, i.e., that the Markov property is satisfied.

In what follows we elaborate on the tuple elements.

States, *S*: the hidden states of the road link status, which describe the safety of the road link. For example, and without loss of generality, two-states can be utilized, whether safe or not. Further in-between states can be added.

Emission Symbols, : the observations from which the hidden states can be deduced. Examples of possible observations are road link congestion rate; road condition metric; road infrastructure type, e.g., number of road link lanes, type side or highway link; or road infrastructure characteristic; e.g., road visibility, etc. A metric utilizing a combination of two or more of these and other observations can also be synthesized.

For illustration, can be made to capture road link congestion, with , representing levels of congestion. If , the congestion can be digitized into levels at , , , all relative to the link's maximum capacity. If we define the random variable as the congestion of road segment, at time, , then, , as a function of road segment congestion, can be defined to be as follows.

It is essential to note, however, that the limits of these levels would need to be normalized to usefully reflect actual links congestion levels.

Transition Probabilities, , is the probability of transition among the two states in the HMM.

Probability of Emission, (or output probabilities), is equal to and given the current state is (safe) or (unsafe) road link, respectively. Extending the illustrative example above, can be calculated if we know the probability distribution of , an empirical distribution interpolated from the sensory data. The empirical distribution can be approximated into the best standard distribution, if possible, taking into consideration the trade-off between accuracy and complexity.

Initial State Distribution, *π*, specifies the initial probability distribution of the states. While typically initialized using a uniform distribution assignment, there are generally no assumptions needed regarding prior distribution.

Three traditional algorithms can be employed to efficiently compute an HMM, namely, the Viterbi algorithm, the Forward-Backward algorithm, and the Expectation-Maximization algorithm [42]. This is particularly advantageous in the computation of link safety for our purposes. Furthermore, once the HMM is determined, the state of the links (as well the state of regions) can be identified and utilized using various services, including route planning, as will be discussed in the next section.

Safety-Based Route Planning

Route planning has become widely used in both personal and commercial use, resulting in an increasing dependence on its reliability. Various applications employ efficient algorithms for route planning [43]. Trip time and cost, e.g., for tolls, have been the typical metrics for route planning applications, but other metrics, however, have been utilized, e.g., for fuel emission/consumption or energy requirements of electric vehicles.

Using the dynamic safety assessment proposed above, it is now possible to route vehicles across cities based on a safety. In this manner, drivers can be directed through routes that minimizes their overall risk in traversing the road network. Meanwhile, enforcement can distribute vehicles across different paths to distribute risk of the network and avoid having critically unsafe links or routes within the network. It is furthermore possible to target auxiliary mechanisms for safety-control across the network by controlling and redirecting traffic based on user driving behavior or in-response to incidental changes in the road network.

An advantage of the assessment core proposed above is that a routing algorithm can be operated directly on its generated values. In what follows we describe a direct application for routes assumed to be traversed shortly after the route have been computed, and that require a traversal time sufficiently less than transition time in the HMM. These assumptions are without loss in generality and can be relaxed with easy modifications to the route planning formulation presented below.

Consider a graph , with comprising nodes (vertices), and comprising the edges in the graph. Nodes represent starting, ending, and midway stops for the vehicle, and our interest is in routing a vehicle from a source node () to a destination node (). Vertices are further identified by numbers, with vertex 0 identifying the source node and identifying possible target destinations. Nonsource nodes make the set .

The set of edges represents the set of links between the nodes. Each edge has an associated traversal cost , which may be either symmetric (or asymmetric (). Herein, this cost can be inversely proportional to a link's safety assessment.

Considering the above, a formulation can be presented as shown in Table 1. The formulation minimizes the expended cost. The first four constraints suffice for a general (uncapacitated) shortest-path problem. The first constraint mandates that a stop along the chosen route is visited only once. The second constraint specifies that, except for the source and the destination, each stop has as many ingoing as outgoing traversals. The third and fourth constraint mandate single departure from the source and a single arrival at the destination.

Table 1
Formulation for the route planning problem.

The above formulation is an integer-programming instance, making the essential route planning problem NP-hard. Meanwhile, when engaging the computation core, as specified by the formulation above, the core is supplied with the updated "costs" and the "capacities" of the edges. This *updating*, while facilitating more relevant solutions, results in solutions that are at best critically stable. Dealing with probabilistic coefficients or constraint thus becomes inevitable when extending considerations for more realistic modelling. Such coefficients and limits, however, only add to the problem complexity.

The application of the road-safety assessment core introduced in Section 4 above overcomes this complexity aspect by simplifying the handling of the probabilistic. Enumerations. Specifically, the formulation can be revised to be applied over the HMM abstraction of the road-network, allow an end-to-end path selection based on safety.

Conclusions

This work illustrates the viability of an economic road safety monitoring and assessment solution through exploiting advances in the Internet of Things (IoT) within the context of smart cities. The introduced architecture facilitates robust and dynamic road safety assessment that complements the Safe System approach motivated by the World Health Organization (WHO), which has been increasingly adopted worldwide. An application of the dynamic assessment framework for route planning is also demonstrated.

Future work involves exploring further applications, especially in the context of raising driver awareness of the road safety conditions during their trips.