Auto Telematics Consulting

Charles Hesketh

August 2021

1 Disclaimer

This project was in collaboration with a major US Property and Casualty Insurance company (henceforth referred to as "the Company"). I lead a team of four, including me, as we consulted on an important project impacting business decisions for their Auto Policies. The specific details and findings of this project are proprietary so this report will only talk about what was done in generalities.

2 Introduction

Many Auto insurers offer a telematics solution for its customers. Drivers in the program sign up and agree to a means of tracking trip information for the purpose of more accurately pricing their policy based on their driving habits. Drivers determined to be lower risk of loss result in discounts. In some programs, riskier drivers are given a surcharge Our analysis of a sample of drivers two main goals. First, we had to determine a "burn in" metric, something that allows us to determine that enough driving had occurred, so the Company would know when a customer had enough driving record to to make a determination on their level of risk. This entails defining driver habit stability using the metrics we had provided. The second goal is to identify any groups of drivers with similar driver habit developments. One example may be a group of drivers that initially exhibits better driving habits, but stabilizes to a much worse driving habit. In our final report, we defined driver habit stability using an algorithm developed by our group, created a recommendation calculating "burn in" metrics, and performed clustering to give further insights into driver behavior.

3 Telematics Background Information

In this section I provide some background information by paraphrasing different papers that give different mathematical approaches to telematics.

3.1 The Use of Smartphones as Telematics Devices

UBI premiums are determined by drivers' scores, which are calculated based on drivers' behaviors on the road. For example, many insurance companies consider harsh braking an excellent indicator of a driver's risk profile because it indicates how observant and aggressive the driver is, and how the driver navigates their trip in the context of other drivers and their surroundings (Skog et al., 2014). Harsh braking is therefore an important factor in computing a driver's score. Other relevant behaviors include speeding, acceleration, time of day, elapsed time, and location (Skog et al., 2014). Various devices can measure these variables, such as devices placed in the on-board diagnostics (OBD) port, smartphones, installed black boxes, and original equipment manufacturer (OEM) probes. However, devices such as OBD outlet devices bring associated hardware, installation, maintenance, and logistics costs (Skog et al., 2014). OBD devices and black boxes, for instance, typically require an additional SIM card and data plan to function. Thus, there is an economic incentive to use devices which are not subject to these costs, such as smartphones.

Skog et al. (2014) researched the strengths and limitations of using a smartphone as the primary telematics device on-board a vehicle. One benefit is that many existing policy holders already own smartphones, which allows them to easily download UBI applications. Smartphones also already come equipped with much of the hardware necessary for recording trip information, such as a GPS for determining location and speed, an accelerometer to detect swerving, and a gyroscope to detect a reference direction and angular rotational velocity (Gyroscope sensors; Skog et al., 2014). However, smartphones can suffer from poor data quality and reliability. For example, smartphones use the GPS sensor to detect speed once per second. The sensor often exhibits an irregular rate of data acquisition, and while the speed data is accurate, it frequently contains undetected outliers. Fortunately, these two issues can be addressed with digital signal processing of the incoming data. Another limitation of smartphones is that they depend on line-of-sight data transfer to and from satellites. Still, the ubiquity of the smartphone allows UBI drivers to use telematics hardware at no additional cost.

3.2 Real-Time Feedback

A key feature of UBI policies is that they enable insurers to collect real-time information about their policyholders. This is advantageous because insurers can better understand their customers' driving behaviors as they operate their vehicles and reward them in a timely manner with premium discounts for safe driving habits. Results from a 2015 experiment with truck drivers from a U.S. transportation company suggest that having an electronic device monitoring driving behaviors in real-time can lead to overall better driving performance (Blader et al., 2015). In this experiment, the electronic on-board recorder (EOBR) system attached to the truck cab makes an audible alarm sound whenever it detects abnormal driving behaviors, thus allowing the driver to correct the behaviors in real-time. Unlike traditional insurance policies which depend primarily on static demographic variables to determine monthly insurance premiums, UBI programs also use telematics information to provide real-time feedback to insurers about customers' driving habits. This provides a more holistic way of profiling customers' risks of being involved in an automobile accident.

3.3 Economic Impact & Consumer Discounts

With telematics, both auto insurers and their clients can reap monetary benefits while making roadways safer. In particular, clients have greater ability to minimize their monthly premium since they can assess real-time information on their driving behavior. For instance, customers can limit distractions (such as answering phone calls while driving), as their device's sensors provide a percentage score of how often they are distracted. Because these drivers have real-time feedback and insights on their driving performance, they can more quickly determine the steps they need to take to improve their driving habits and increase the discount on their monthly premium. As a result, auto insurers have greater opportunities to establish strong customer loyalty since their customers will likely appreciate the reduction in their premiums. In fact, a 2017 study involving about 95,000 traditional insurance (non-UBI) customers and 40,000 UBI customers showed that UBI customers are 9% more likely than their traditional insurance counterparts to renew their auto insurance policy with their current insurer. In addition, the UBI cohort exhibited faster improvement in driving behavior over time, justifying "adoption of UBI as a way to improve profits, even after considering the costs of the program and the discounts provided" (Soleymanian et al., 2019). Most importantly, the UBI model promotes a safer driving environment for motorists by reinforcing responsible driving behaviors. UBI programs also provide benefits to society, including "reduced accident frequency and severity, reduced accident response time, and ease of establishing fault in settling claims" (Husnjak, Peraković, Forenbacher, & Mumdziev, 2015). In addition, they can reduce customers' driving, which means added environmental benefits in terms of pollution and fuel consumption.

3.4 Estimating the Necessary Amount of Driving Data for Assessing Driving Behavior

When analyzing drivers, insurers may want to understand the consistency of their insureds' driving behaviors. One way to measure the stability of a driver's behavior on the road is to see how much their overall driving performance fluctuates over a certain period of time, number of trips taken, or number of miles driven. However, quantifying the exact number of trips/miles needed for performance score to stabilize for a large, diverse population is complex. One reason is that road networks in one geographic area may be different from road networks in another area, such that a driver's behaviors and needs may vary significantly based on the location. For instance, given similar driving performance, a driver living in the urbanized metro area of New York City may have a lower mileage per trip than another driver in rural farmland of South Dakota, as the roads in New York City are much more condensed (shorter point-to-point destinations). Also, the total number of trips a driver may take in New York City may be less than that for a driver living in Keystone, South Dakota, as New York City provides other modes of public transportation (Stavrakaki et al., 2020).

Stavrakaki et al. (2020) attempted to create a methodological framework for estimating how much information is required to understand an individual's driving behavior. They focused on determining a minimum amount of driving time needed to group drivers' aggressiveness. A driver associated with high aggressiveness was an indicator of lower scores and faster stabilization, while low aggressiveness tended to have higher scores but would take much longer for their scores to stabilize. OSeven Telematics' mobile application used the driver's cell phone accelerometer, magnetometer, gyroscope, and GPS to anonymously record data points at a 1 Hz frequency without the need for user interaction. In total, 21,610 trips were recorded from 68 drivers and processed by OSeven Telematics prior to analysis by the authors.

The authors observe that known risks for more serious accidents include: total distance driven, road network type (city vs. rural), "risky hours", unfamiliar environment, weather, vehicle type, not wearing a seat belt, phone usage, speed, and harsh breaking/acceleration/cornering. While not all of these factors can be recorded, a smartphone can measure several of these variables to provide a reasonable risk assessment. Stavrakaki et al. (2020) used trips of length 5, 10, and 20 minute categories, and frequency of each category was used to understand how driving evolved over time. Moving average, volatility, heart charts, and principles of Shewhart charts were used to develop these convergence criteria (directly from the paper):

- "Moving average is within the range mean $\pm$ 1 standard deviation.
- For five consecutive trips, the percent change (in absolute terms) between successive values of the moving average is less than or equal to 1.5%.
- The value of the moving average in the corresponding trip is a local extreme (this criterion ensures that the neighboring values of the moving average are smaller or larger than the selected one, and therefore, it does not belong to a sequence of points that have a particular trend e.g., ascending or descending)."

By analyzing each driver's cumulative sum and volatility, the researchers were able to determine each individual driver's stability. Let k be the metric observed, i be the driver, and t be the number of trips. Then, the Gain/Loss Ratio (needed for volatility) was calculated by the authors as

$$r_{t,i} = ln\left(\frac{k_{t,i}}{k_{t-1,i}}\right). \tag{1}$$

Importantly, this will return a negative value when the metric is improving and a positive value when the metric is worsening. With this, the authors calculated volatility as

$$Volatility = \sqrt{\frac{\sum_{t=1}^{n}(r_{t,i} - \bar{r_i})^2}{n-1}}. \tag{2}$$

Since each driver had a different number of trips, the authors used a moving window of 20 trips for the calculations above.

This study analyzed driving behaviors using metrics such as "number of harsh acceleration and braking events, the duration of mobile usage while driving and the percentage of time driving over the speed limits." Results suggested that it is "extremely difficult to identify the exact time point where a driver's behavior converges when the trips being studied do not have a similar duration" Stavrakaki et al. (2020). Realistically, not all drivers' scores may converge to certain values within the thirteen month time-span (January 2020 to February 2021) of our data set, due to possible outliers caused by external situations such as the COVID-19 pandemic.

4 Complications

The Company did not provide raw data to our group, instead they provided values based on a proprietary transformation that we were not privy to. This made it impossible to calculate many typical statistics that we would have liked to utilize such as variance. There was a change in the calculation of this transformation that caused "lost data" previous to this switch. In addition the data had new user entering the data set prior to and after COVID where lockdowns changed how often many drivers drove.

5 Results

Given our data limitations, we explored different ways to approach the problem. Instead of the data being given in discrete trips, or scores, we were given a daily snapshot for each driver, which did not allow us to directly implement methods used in Stavrakaki et al. Given our inability to directly calculate co-variance we were unable to determine stability based on variance. Given this restriction we created an algorithm that could compare the fluctuation of scores over time based on ranges. To verify our approach we gained inspiration from the stock market and fit a GARCH model. This verified our initial findings and supported the potential usefulness of our model.

In addition we provided an analysis using clustering algorithms that provided insights into the features of drivers with similar habits. This analysis allowed the Company to have a deeper understanding of their customer base. It allowed the company to make inferences on its customers based on geographical, urban vs rural, and future analysis with features that were not provided to us.

We detailed everything in a report that was over 20 page and provided it to the Company. We created a two page report geared toward C-Level execs with the goal of explaining the results within the first half page and the rest of the document was a high level summary of our process. The team provided several coding files to the company to be used by their analytic team in the future. This all culminated in an hour long presentation to high level management within the Organization where we detailed our process, explained our findings, and defended our results. Everything was well received by the Company and they were very happy with our results.

6 References

Blader, S., Gartenberg, C., & Prat, A. (2015). The contingent effect of management practices. *SSRN Electronic Journal*. doi: 10.2139/ssrn.2594258

How does a gyroscope sensor work in your smartphone? (n.d.). https://www.techaheadcorp.com/knowledge-center/how-gyroscope-sensor-work-in-smartphone. (Accessed: 7-March-2021)

Husnjak, S., Peraković, D., Forenbacher, I., & Mumdziev, M. (2015). Telematics system in usage based motor insurance. *Procedia Engineering*, *100*, 816-825. https://www.sciencedirect.com/science/article/pii/S1877705815004634. doi: https://doi.org/10.1016/j.proeng.2015.01.436

Skog, I., Bonawiede, F., Welch, R., Ohlsson, J., Händel, P., Wahlström, J., & Ohlsson, M. (2014). Insurance telematics: Opportunities and challenges with the smartphone solution. *Submitted to IEEE Intelligent Transportation Systems Magazine*, *6*. doi: 10.1109/MITS.2014.2343262

Soleymanian, M., Weinberg, C., & Zhu, T. (2019). Sensor data and behavioral tracking: Does usage-based auto insurance benefit drivers? *Marketing Science*, *38*. doi: 10.1287/mksc.2018.1126

Stavrakaki, A. M., Tselentis, D. I., Barmpounakis, E., Vlahogianni, E. I., & Yannis, G. (2020). Estimating the necessary amount of driving data for assessing driving behavior. *Sensors (Basel, Switzerland) 20*, *(9):2600*. doi: 10.3390/s20092600