Accident Prevention with Predictive Instantaneous Crowdsourcing

John Joon Young Chung
jjyc@umich.edu
University of Michigan, Ann Arbor

Fuhu Xiao
fuhu@umich.edu
University of Michigan, Ann Arbor

Nicholas Recker
nrecker@umich.edu
University of Michigan, Ann Arbor

Kammeran Barnes
kameranb@umich.edu
University of Michigan, Ann Arbor

Nikola Banovic
nbanovic@umich.edu
University of Michigan, Ann Arbor

Walter S. Lasecki
wlasecki@umich.edu
University of Michigan, Ann Arbor

ABSTRACT

Training machine learning algorithms to avoid traffic accidents can be challenging because the rare occurrence of such events leads to the insufficiency of training data. We introduce the idea of applying instantaneous crowdsourcing to augment autonomous vehicles with collective human cognitive capability within super-human reaction time. However, because the instantaneous crowdsourcing system must prefetch possible futures in order to generate tasks, in complex real-world problems we would need to hire implausibly many workers to support this approach. In this work, we propose that predicting dangerous futures from crowd-worker input can help resolve this problem. In a formative study to inform the design of crowd prediction workflows, we found there are two main challenges: 1) false positives, which can initiate instantaneous crowdsourcing more than necessary, and 2) handling a large number of futures with multiple candidate objects in the scene.

CHI'19 Workshop on "Looking into the Future: Weaving the Threads of Vehicle Automation", May 2019, Glasgow, UK

Figure 1: Screenshot of the Crowdsourcing Task Interface. If a crowd worker thinks that a vehicle in the video scene will cause an accident, she can annotate the vehicle by clicking and tracking it.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**.

KEYWORDS

autonomous vehicles, real-time crowdsourcing, instantaneous crowdsourcing, prediction

INTRODUCTION

As autonomous driving technology rapidly develops, there are increasing concerns about whether dangerous accidents can be avoided in complex situations [1]. Current machine learning models are limited at coping with dangerous events because such events are rare, so there is not enough data to train a reliable model [7]. Additionally, such events are also difficult for humans to avoid, because events can happen faster than human cognitive processing time.

We propose a model in which the system can switch to using human oversight [5] enabled by instantaneous crowdsourcing [13] (which provides human capability at speeds faster than cognitive processing time via collective action) when a machine learning models are not capable of handling a scenario encountered [11]. Real-time crowdsourcing has solved many computing problems that require agile reaction of a system and cannot be solved solely by a machine [2, 8–10]. Instantaneous crowdsourcing greatly accelerates real-time crowdsourcing to provide answers within a couple milliseconds by prefetching tasks based on possible futures and using pre-collected task results "instantaneously" (at system lookup speeds) when a target state is detected [13].

However, in real-world problems like avoiding vehicle-to-vehicle collisions, the number of futures can be far too large to apply instantaneous crowdsourcing. For example, a vehicle in front of us can move in any direction at any speed and with multiple dynamic objects in the scene, the number of possible futures can increase exponentially. Therefore, prefetching tasks for all different possible futures would require a tremendous number of workers. It would increase the cost and possibly delay the instantaneous crowdsourcing if we cannot recruit enough crowd workers in time.

To overcome these limitations and enable the use of instantaneous crowdsourcing in autonomous driving settings, we propose a crowd workflow that predicts the most probable futures. By leveraging the prediction capabilities of humans shown in situational awareness [4] and defensive driving [3], the workflow reduces the number of futures to those that are most probable and relevant. With the minimized number of futures, we would be able to use instantaneous crowdsourcing in real-world problems more economically.

However, not as much is known about human prediction capability as it relates to instantaneous crowdsourcing. To inform the design of a prediction workflow, we first conducted a formative study, examining how crowd workers make predictions in different accident situations. From the result, we

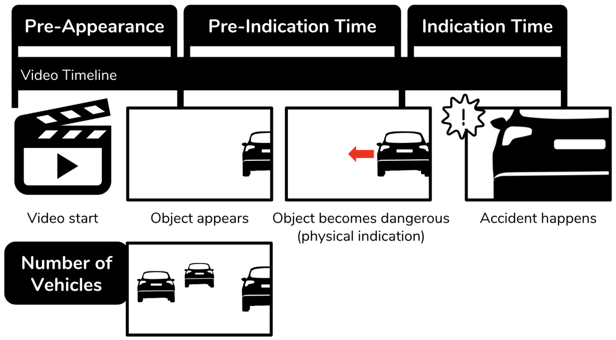

Figure 2: Independent variables in accident simulation videos

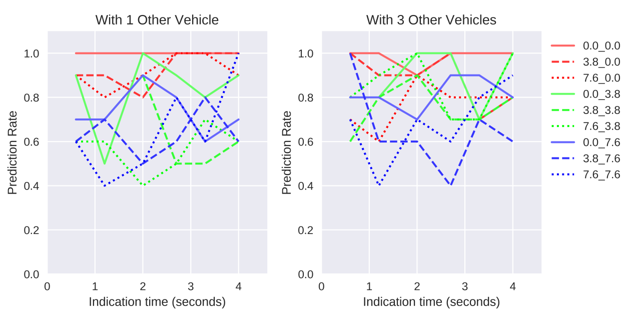

Figure 3: Rate of crowd workers capturing the accident with the first prediction, after the indication, in videos with an accident. The first value in the legend indicates *Pre-appearance time* and the second value is *Pre-indication time*.

found two major challenges in designing the prediction workflow: 1) false positives and 2) workers handling a large number of futures from many candidate objects.

FORMATIVE STUDY

To design a workflow that predicts traffic accidents, we conducted a study to understand to what extent people are capable of making predictions in different conditions. In the task, crowd workers watched videos with or without an accident, simulated on Unity3D. Workers annotated vehicles that they thought would cause an accident in the near future. Lest workers think that an accident exists in all videos, we gave them two videos as tutorials: one with an accident and one without. For conditions of accidents, we came up with 4 independent variables to vary (Fig. 2):

- *Pre-appearance time*: The time between the start of the video and the appearance of the accident vehicle, which was chosen from 0, 3.8, and 7.6 seconds.
- *Pre-indication time*: The time between the appearance of the accident vehicle and it giving a clear indication of danger, which was chosen from 0, 3.8, and 7.6 seconds.
- *Indication time*: The time between the accident vehicle giving a clear indication of danger and the accident, which was chosen from 0.6, 1.2, 2.0, 2.7, 3.3, and 4.0 seconds
- *Vehicle num*: The number of vehicles in the video, which was chosen from 1 and 3 vehicles

We manipulated the first three variables to see the effects of time on prediction. In *Pre-indication time* and *Indication time*, we defined the physical indication of danger as the moment when the vehicle with the camera and the accident vehicle are expected to collide in near future if they maintained their current physical dynamics. In the videos we generated, the indication was the accident vehicle cutting into the lane. We also tried to see the effects of scene complexity by manipulating the number of vehicles.

To measure the prediction capability, we used three metrics for videos with an accident: 1) **prediction time**, the time between a crowd worker making the first correct prediction and the accident and 2) **prediction rate**, how many crowd workers were capable of predicting the accident correctly after the indication. Here, we only considered predictions made after the indication as correct predictions, because predictions without any clues have a lower probability of turning out to be an accident. These might unnecessarily start the instantaneous crowdsourcing workflow in a real-deployment. With this consideration, for videos without an accident, we used 3) **false positive rate**, how many crowd workers made false positives before the indication, as the last metric.

RESULTS

We analyzed three metrics with linear regression. For the prediction rate (Fig. 3), across all videos with an accident, 78.7% of crowd workers could capture the accident. The prediction rate was negatively correlated with *Pre-appearance time* (coeff=-.018, p<0.001) and *Pre-indication time* (coeff=-0.029, p<0.001; $R^2 = 0.395$). It might have been because crowd workers could make false positives before the indication

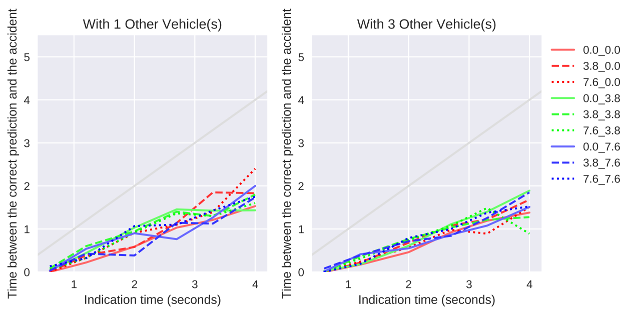

Figure 4: Mean time from correct prediction to accident for different variables, for annotations that were made after indications. The first value in the legend indicates *Pre-appearance time* and the second value is *Pre-indication time*.

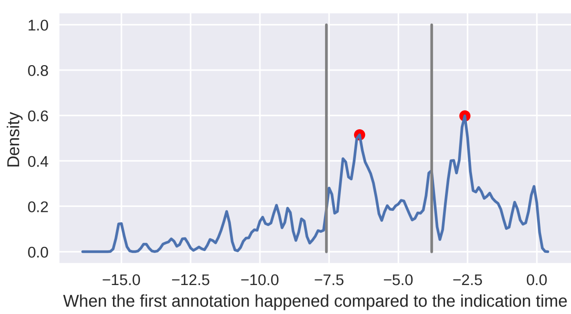

Figure 5: Smoothed histogram of the first annotation time for workers who made the annotation before the indication time. Time 0 is the indication time and negative values mean that the first annotations were made before the indication. Red dots indicate the peak calculated with topographic prominence of 0.25

with a longer *Pre-appearance time* or *Pre-indication time*. Including these false positives, we observed that crowd workers' performance did not depend on the *Indication time*, or how quickly the object is moving. For false positive rate in videos without an accident, 30% of crowd workers made a false positive prediction before an indication of danger existed. *Pre-appearance time* was positively correlated with false positive rate (coeff=0.025, $p<0.05$; $R^2 = 0.548$).

For the prediction time, we found that 20.7% of crowd workers annotated the accident vehicle even before the indication, which we define as premature predictions. Because we considered this behavior to be a false positive and wanted to understand how quickly crowds react to the indication, we conducted an analysis without premature predictions (Fig. 4). We found that prediction time was positively correlated with the *Indication time* (coeff=0.47, $p<0.001$), and negatively correlated with *Vehicle num* (coeff=-0.078, $p<0.001$; $R^2 = 0.531$). This suggests that crowd workers might have reacted to the indication when making the prediction. Across videos, we observed that the crowd workers made the first prediction 1.49 seconds (σ=0.76) after the indication, while the best performing crowd workers recording 0.90 seconds (σ=0.42).

However, this performance would only be retrieved when there is no worker who makes a premature prediction. In order to investigate the behaviors of premature predictions, we ran an additional analysis on when crowd workers made the first premature annotation, for both accidental and non-accidental videos. We observed that a relatively large number of crowd workers started annotation right after the appearance of the accident vehicle while other workers annotating regardless of time.

CONCLUSIONS AND FUTURE WORK

We claim that false positives can be a challenge in the prediction workflow for instantaneous crowdsourcing. We observed that people's interpretations of the danger vary, which might be due to different perspectives that people have for the danger [6, 12]. The premature prediction is one example of the false positives that can be caused by different risk perception. Because the purpose of the prediction workflow is to initiate an instantaneous crowdsourcing workflow only when an accident is expected to happen with a high probability, we need to design a workflow that suppresses these false positives.

Additionally, dealing with a large number of potential futures with many variable objects can be another challenge for the prediction workflow. For example, with more vehicles, we observed the delay of crowd predictions. It might signal human's limited capability in processing a large number of futures. It leads to the need for distributing prediction tasks on multiple vehicles to different workers.

ACKNOWLEDGEMENT

This research was supported by U.S. DOT Center for Connected and Automated Transportation (CCAT).

REFERENCES

[1] Pavlo Bazilinksyy, Miltos Kyriakidis, and Joost de Winter. 2015. An international crowdsourcing study into people's statements on fully automated driving. (2015).

[2] Michael S Bernstein, Joel Brandt, Robert C Miller, and David R Karger. 2011. Crowds in two seconds: Enabling realtime crowd-powered interfaces. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 33–42.

[3] DriversEducationUSA.com. 2019. What is Defensive Driving? https://www.driverseducationusa.com/resources/what-is-defensive-driving/. Accessed: 2019-01-31.

[4] Mica R Endsley. 1995. Measurement of situation awareness in dynamic systems. *Human factors* 37, 1 (1995), 65–84.

[5] Sai R Gouravajhala, Jinyeong Yim, Karthik Desingh, Yanda Huang, Odest Chadwicke Jenkins, and Walter S Lasecki. 2018. EURECA: Enhanced Understanding of Real Environments via Crowd Assistance. In *Sixth AAAI Conference on Human Computation and Crowdsourcing*.

[6] John A Groeger and PR Chapman. 1996. Judgement of traffic scenes: The role of danger and difficulty. *Applied Cognitive Psychology* 10, 4 (1996), 349–364.

[7] Nidhi Kalra and Susan M Paddock. 2016. Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? *Transportation Research Part A: Policy and Practice* 94 (2016), 182–193.

[8] Walter S Lasecki, Mitchell Gordon, Danai Koutra, Malte F Jung, Steven P Dow, and Jeffrey P Bigham. 2014. Glance: Rapidly coding behavioral video with the crowd. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM, 551–562.

[9] Walter S Lasecki, Christopher Miller, Adam Sadilek, Andrew Abumoussa, Donato Borrello, Raja Kushalnagar, and Jeffrey P Bigham. 2012. Real-time captioning by groups of non-experts. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*. ACM, 23–34.

[10] Walter S Lasecki, Kyle I Murray, Samuel White, Robert C Miller, and Jeffrey P Bigham. 2011. Real-time crowd control of existing interfaces. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 23–32.

[11] Christopher H Lin, Mausam, and Daniel S Weld. 2012. Dynamically Switching between Synergistic Workflows for Crowdsourcing. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*.

[12] Ingunn Olea Lund and Torbjörn Rundmo. 2009. Cross-cultural comparisons of traffic safety, risk perception, attitudes and behaviour. *Safety Science* 47, 4 (2009), 547–553.

[13] Alan Lundgard, Yiwei Yang, Maya L Foster, and Walter S Lasecki. 2018. Bolt: Instantaneous Crowdsourcing via Just-in-Time Training. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 467.