"How autonomous vehicles can talk to one another"

Monday, June 16, 2025

Guests:

YONG LIU
Professor of Electrical and Computer Engineering,
New York University

Guest Host:

DARRELL WEST,
Co-host, The TechTank Podcast;
Senior Fellow,
Center for Technology Innovation,
Governance Studies,
The Brookings Institution

[music]

CO-HOST NICOL TURNER LEE [00:00:00] You're listening to Tech Tank, a bi-weekly podcast from the Brookings Institution, exploring the most consequential technology

issues of our time. From racial bias and algorithms to the future of work, Tech Tank takes big ideas and makes them accessible.

CO-HOST DARRELL WEST [00:00:27] Thanks for joining our Brookings Tech Tank podcast. I'm Darrell West, a senior fellow in the Center for Technology Innovation at the Brookings Institution. It is an eventful time in the development of autonomous vehicles. Taxis are appearing on the streets of San Francisco, Phoenix, and Austin, among other places, and there are self-driving trucks in Texas. Experts anticipate these cars and trucks will spread to other cities in the very near future. At Brookings, we have published a new paper entitled "How autonomous vehicles could change cities". It talks about how the expansion of self-driving vehicles could reduce accidents and create opportunities for cities to redesign streets and urban areas for biking, pedestrians, and automated delivery systems. You can read more about that paper at brickings.edu. But there also are interesting questions regarding how these vehicles can navigate roads and do a better job communicating with one another. A new research project from New York University investigates ways self-driving vehicles can improve their efficiency and navigational systems by creating social networks of autonomous vehicles. To help us understand this improvement, we are pleased to be joined by one of the authors of that study. He is Professor Yong Liu, and he is a professor of electrical and computer engineering at New York university. He and I will discuss what this new system is and how it can improve reliability and efficiency in autonomous vehicles. Yong, welcome to our Brookings tech tank podcast.

GUEST YONG LIU [00:02:05] Thank you, Darrell, for having me. My pleasure to share with everybody about our recent work.

CO-HOST DARRELL WEST [00:02:12] So let's jump in and hear about that project. So I know you've looked at how autonomous vehicles can use AI to share information on the road and improve navigation. So tell us what that system is and how it operates.

GUEST YONG LIU [00:02:28] So traditionally, when people talk about vehicle networks, they're talking about using communication between vehicles to kind of collect data, because each vehicle have lots of sensors. They collect data from the surrounding environments as well as data from inside of the vehicle. So there are different ways to do that. But now we are talking about something a little bit different, because as you can imagine, that the data collected from each vehicle, There are some kind of a... Privacy

issues, because the owners of the vehicle may or may not want to share the data with other vehicles or with the public. But at the same time, as I said, those data are very critical, very valuable, to make the whole vehicle system better in terms of come up with better scheme to do the routing and come up better scheme to do object detection and so on and so forth. So how do we kind of... Utilize those data to derive intelligence, at the same time preserve the privacy of individual owners of the vehicles. So that's a challenge we are facing. So the system we are developing is exactly to address this issue. We resort to this new paradigm for learning, it's called federated learning. So, this new paradigm allows agents to participate in collective training. Without sharing their private data. So in this vehicular kind of a system in particular, what we are looking at here is each vehicle still collecting data, but they don't share data with other vehicles or with any other central server. So what are they going to do? They're going to use some kind of local computation power to train some model. So the model can be for object detection, that can be routing, or... Whatever you can think about. So they're going to train some intelligent local models. Then they are going to share the local models with other vehicles without sharing the data. But now, of course, vehicles are moving. So how do you enable sharing among vehicles is another challenge, what we call the communication challenge, which I have to address. So I'm part of the research center at NYU is called NYU Wireless. We're looking at 5G, 6G, and also part of the Center for Advanced Technology and Telecommunication at NYU. Again, we're looking on networking issues for all kinds of systems. Going back to the vehicle system, so there are different modes of communication between agents. So one mode is like upload your model to a central server using long-range communication, let's say 5G connection. And like what you are using on your cell phone but the problem with that approach here is that the server might be far away from you and you have to use a long run to communication to do that and if you're talking about deep learning models those models tend to be huge so which means you need to also upload large amount of parameters to the server and another issue with that is the server, you're thinking about... You know, hundreds, thousands of vehicles, they all want to upload to the same server, the server quickly become bottleneck. So both in terms of performance and reliability. So what we are looking at is a solution, what we call decentralized federated learning. So in this example here, you want to get rid of the central server. You want to rid of that single point of failure. So what it does is it allow vehicles to talk to each other directly. Can think about, as Darrell mentioned, it's like a build a social network among vehicles so that they can do vehicle-to-vehicle communication directly. So luckily, with the new advance in 5G, 6G, it's possible to do this short-range device-to device vehicle- to vehicle communication at a high speed and also at a low latency. So that is, whenever two cars, two vehicles get close to each other, for example, they're stopping at the same traffic light, waiting for the the light, they can in a close proximity, they can exchange data at high speed and low latency. That's where they're going to do what we call local

model exchange. So they're gonna get each other's model and then they try to learn from each other, try to come up with a better way of doing, for example, navigation.

CO-HOST DARRELL WEST [00:07:02] OK, now that's very helpful to get a sense of how that operates. And as I understand it, the vehicles will be able to share information about navigational obstacles. They may be encountering traffic patterns, road conditions, traffic signals, and the like, and then adjust their navigational capabilities based on that information. And it seems like you ran a simulation with 100 virtual self-driving cars in Manhattan. And each vehicle had AI models built into it that updated on a regular basis and then shared the information. Tell us a little bit about this simulation and how that allowed you to devise a system that would improve navigational systems.

GUEST YONG LIU [00:07:44] So one important part of the kind of a navigation is to do object detection. So when cars have all kinds of sensors, camera, ladder, and they collect the data from the environment. And it should be quickly recognized what kind of objects is passing by, so that they can make the navigation, first of all, to be safe. And also, can figure out what is the congestion level right now. Random. Do the right strategy. So for our simulation we start with the basic task, that's object detection. And so we use simulation, we have like 100 cars in the Manhattan street kind of network, and we use some kind of a mobility model, so which means whenever a car drives to an intersection and there's a certain probability that goes straight, turn left, turn right, which means mobility is pretty random. Under this random mobility of all the vehicles and all the vehicles going to train some object detection model and we want to look at how fast the model converge, what is the final converge accuracy in terms of object detection. So what we are comparing with is the traditional centralized training by assuming all the data upload to the central server and the central server train object detection So obviously, you can think that will be the upper bound. That's the best you can do in terms of convergence speed and the final accuracy. And so indeed, in our simulation, what we propose that is using Kaishen on the vehicle to help the convergence can approach the upper bond. That is a centralized training. But we don't have a central server. So we rely on device to device communication. We rely on computation power on individual vehicle instead of data centers, server data centers to achieve more or less the same accuracy.

CO-HOST DARRELL WEST [00:09:45] So I understand there's no central server. So this is a decentralized communications network. So there's not one central location where all the information is stored. And I'm sure that helps with any potential

catastrophic data breaches. So it sounds like that helps a lot on the cybersecurity front. But what about the privacy issue? People always are worried about information being shared with other people. How does it help on cybersecurity? And then what can you do to protect privacy of people's information.

GUEST YONG LIU [00:10:17] So indeed, our starting point is to protect the privacy of users. So as I mentioned, the users data never get out of their car, so which means we never upload user data anywhere to the server. So by the way, vehicles do have to exchange the model parameters, for example. So there are some studies saying that even though you don't share the data itself, but if you share the model parameters with other users, they may do some reverse engineering and try to infer your private data. So for that, I guess there are also solutions. People work on different ways to do the secure model aggregation, and they use different techniques such as differential privacy so that you can get the same aggregation result without disclosing individual users' model updates. So those are the solutions we can use. But I said, the bottom line here is that... Our research is indeed trying to protect user privacy and by using this decentralized federal learning framework. But then, of course, there are some related issues about, you know, what if some vehicles are malicious? So, for example, they inject the bad model updates, try to collect the system. Again, that's the common issue with any distributed system, large-scale peer-to-peer A lot of research has gone into that in different domains and some of the solutions are ready to be adopted in our system as well to deal with what we call the pollution attack on the system.

CO-HOST DARRELL WEST [00:12:00] So how scalable are the systems that you're proposing? You obviously were basing your work on this Manhattan simulation. Could that work for a city as a whole?

GUEST YONG LIU [00:12:11] We have been working on this large-scale peer-to-peer system over many years. We started with peer- to-pear file distribution. Some of you may have heard about the name, about Kazaa, BitTorrent, so on and so forth. So one unique advantage of this system is self-scalable. So which means when you add more peers into the system, they bring resources to the system and make the system self-scalable. And another important example for peer-to-peer system is Bitcoin, obviously. So there is a huge system and lots of coin miners join the system. And the system indeed relies on this large-scale distributed computation to achieve kind of trust, so nobody can control the system and they don't have centralized authority. To trust so they have that kind of a bitcoin system. Coming back to our proposed, you can call it the

peer-to-peer motion learning system, peer- to-pear intelligence system. And one interesting, very attractive feature is, again, the self scalability. So as you put more cards into the system, they are going to bring more data, obviously. They're also going to bring more computation power. So we assume each car has some computation power available. So now finally about the communication side. So we are relying on this vehicle-to-vehicle communication, so which means whenever two vehicles meet each other, they are going to do high bandwidth, low latency communication, and by itself, it's also scalable. So which means, you know, whenever you have more cars, they have more chances to talk to other cars, right? So obviously, you now, from the whole network point of view, you need to figure out a way to organize a large number of vehicles and different topologies you can build among them. That's part of our research. So essentially, how are you going to scale up the system? But in general, the more vehicles join the system, you are going to achieve better efficiency.

CO-HOST DARRELL WEST [00:14:21] So it seems like the next step would be real-world testing of this system, and I'm just wondering are there any plans for that?

GUEST YONG LIU [00:14:30] Exactly. So we currently, you know, we just published a paper early this year in the Tribal AI conference. You know, as you mentioned, the current results is just purely based on simulation. Of course, we want to do much more than that. So I'm talking to some experts working on in a transportation system and to say whether we can do some more realistic evaluation either using more realistic models or realistic traffic conditions. And eventually we may want to do some kind of demos on the Ray

CO-HOST DARRELL WEST [00:15:03] So I'm curious how you overcome different brands of cars and trucks. Are there any communications issues across brands? Do Ford models communicate clearly with General Motors vehicles?

GUEST YONG LIU [00:15:17] Right, so in terms of the communication, I think as long as they follow the standards and they should be able to talk to each other, it's just like if you have different cell phones from different manufacturers, as long they follow same standard, they should able to to talk with each other. And on the model side, of course, they need to share some kind of common architecture for the models so that they can communicate, they can transfer their knowledge, right? And in terms of data, again, they have different formats for the data they collected. Obviously, they need to have some

kind of, you can say, converter or riper so that they can process the data. Finally, the data should be utilized to train a common motion learning model architecture.

CO-HOST DARRELL WEST [00:16:09] So you mentioned the issue of standards, and I would imagine that would be an important part of this system as well. How can we enable communications with traffic signals and other connected devices? Do we need more technical standards in this area in order to facilitate autonomous vehicles? Yeah, definitely.

GUEST YONG LIU [00:16:28] As I mentioned, lots of research have gone into the vehicle network. In the early days, they were focused on how to enable data sharing among vehicles. There are some kind of lots of efforts in the standard body as well to enable that. So we are going to leverage on what has been established there for our research. But on top of that, additional challenges as well, because we are not just talking about it. Exchanging data between vehicles, we are talking about due to the kind of intelligence or model aggregation and of course there should be additional standards in terms of how the data will be represented and how the knowledge will be representative, how you can do the sharing and those are the things that details need to be figured out before this becomes kind of a reality.

CO-HOST DARRELL WEST [00:17:22] So I'm curious, how optimistic are you about the future of autonomous vehicles? Do you think that we are nearing a point where they're going to become more and more present, at least in large American cities?

GUEST YONG LIU [00:17:34] Myself, again, I want to say, first of all, I'm not an expert really on the transportation side. I'm more from networking and the disability-affected learning side, but myself, I kind of have, I think it's very promising. And I think the different parts of the world, and they're already doing some experiments about autonomous vehicle. And they are only on the road already. So, overall, I think I'm pretty, I'm kind of a... I'm pretty sure it's going to happen and also what we are working on is actually hopefully will contribute to make it more efficient and more reliable.

CO-HOST DARRELL WEST [00:18:13] Well, thank you very much. I want to thank Yong Liu for sharing his thoughts with us You can read his articles on the New York University website. And here at Brookings, we write regularly about autonomous

vehicles and digital technology. And you can find more information on our Brookings Tech Tank blog located at Brooking.edu. Thank you very, much for tuning in today.

GUEST YONG LIU [00:18:34] Yes, thank you for having me here.

CO-HOST NICOL TURNER LEE [00:18:45] Thank you for listening to Tech Tank, a series of roundtable discussions and interviews with technology experts and policymakers. For more conversations like this, subscribe to the podcast and sign up to receive the Tech Tank newsletter for more research and analysis from the Center for Technology Innovation at Brookings.