MINNESOTA GO

Planning Minnesota's
Transportation Future

BIG DATA AND CYBERSECURITY TREND ANALYSIS

CONTENTS

INTRODUCTION

Many of our daily activities produce data. Activities such as using social media, mapping a commute, and carrying a smartphone generate data. This data is collected, analyzed, and increasingly being used to improve transportation networks across both private and public sectors. Data is getting larger and more complex as our work becomes increasingly connected by technology. These datasets are sometimes referred to as big data, defined as data gathered from devices like smartphones and services like online shopping.[1]

The breadth of big data creates opportunities to reimagine how we live. Big data helps the healthcare system understand how treatments and procedures impact patient populations. Retailers use big data to better target customers and to suggest products to consumers. And the Minnesota Department of Transportation uses big data to maintain and improve our roads and bridges and understand traffic flows and transportation trends.

Cyber security is vital for protecting big data. The United States Department of Homeland Security defines cyber security as "The art of protecting networks, devices, and data from unauthorized access or criminal use…."[2] Security concerns multiply as the abundance of collectible data grows—generated from smartphones, Wi-Fi-enabled devices, and automated vehicles. Cyber security is no longer just an ad blocker on your computer. Cyber security is now concerned with smart vacuums and coffee makers. While it is handy to brew coffee remotely, the coffee maker's Wi-Fi capabilities make it vulnerable to cyber-attacks. Likewise, connected traffic signals or dynamic message signs might also be vulnerable.

Protecting these systems will save individuals and institutions from erased databases, corrupted systems, and files, automated vehicles being hacked and taken control of, and stolen personal information.

BIG DATA IN TRANSPORTATION

IMPROVED PLANNING AND PROCESS

The Minnesota Department of Transportation (MnDOT) uses big data to improve roadways. MnDOT has contracted with the company Streetlight to access data to analyze traffic counts and travel patterns. Streetlight uses anonymized data collected from smartphones and GPS navigation devices to remove any information that would identify individuals. MnDOT can analyze the anonymized data to examine how people walking navigate a highway overpass, how drivers use side streets to get around closed roads, and how freight traffic enters and exits a town. [3] Without GPS and cellphone data, MnDOT staff would have had to go to these locations, manually count traffic, and spend substantial time and money.

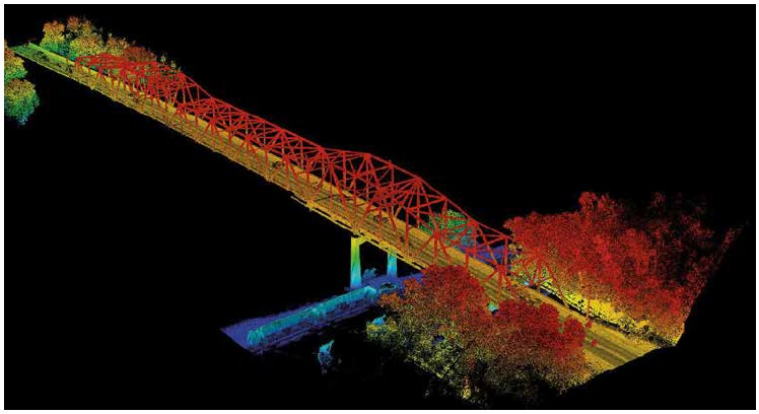
Smartphone data used to track car movement is an example of a transportation-related application of the Internet of Things (IoT). Every Wi-Fi enabled device, from toasters to thermostats, allows these devices to send

[1] Michael Mattioli, "Disclosing Big Data," *Minnesota Law Review* (November 2014): 539-40.
[2] United States Department of Homeland Security, "Security Tips (ST04-001): What is Cyber security," National Cyber Awareness System, https://www.us-cert.gov/ncas/tips/ST04-001 (accessed April 6, 2020).
[3] Streetlight Data, "How it Works," https://www.streetlightdata.com/ (accessed February 6, 2020).

and receive data. The data is collected and summed feeding big data applications. These devices, including smartphones, often have limited or no defenses against cyber security attacks. Transportation agencies including MnDOT are considering how to safely integrate IoT devices in the transportation system.

MnDOT also uses mobile light detection and ranging (LiDAR) technologies to collect surveying and asset data such as signals, lights, bridges, interchanges and weigh stations. Figure 1 shows this data at work.[4]

Figure 1: LiDAR Imagery of Highway 63 Bridge in Red Wing



By capturing this information in a point cloud,[5] technicians can extract the data and create a 3-D image of the asset—there's less need to examine the actual guard rail when there's a digital, 3-D copy. These LiDAR point clouds have varying accuracies. MnDOT's mobile LiDAR data range from an absolute vertical accuracy of about half an inch for surveying to as much as three feet for asset collection.[6,7] This allows staff to estimate project costs and constraints remotely and determine fixes to damaged assets without physically standing at the location. Before LiDAR data collection of big data, all of these tasks were manual.

[4] Minnesota Department of Transportation, "Putting Research into Practice: Using Mobile Mapping to Inventory Barriers," Research Services & Library, 2014.

[5] A Point Cloud is "a collection of three-dimensional coordinates (latitude, longitude, and height) that correspond to a particular point on the Earth's surface from which a laser pulse was reflected. The point clouds are used to generate other geospatial products, such as digital elevation models, canopy models, building models, and contours." National Oceanic and Atmospheric Administration, "What is LIDAR?" Under Ocean Facts, https://oceanservice.noaa.gov/facts/lidar.html (accessed April 6, 2020).

[6] Minnesota Department of Transportation, "Putting Research into Practice: Using Mobile Mapping to Inventory Barriers," Research Services & Library, 2014.

[7] The asset LiDAR data is collected using a GPS mounted on the vehicle and the location of the data is usually off around 3 feet in absolute accuracy meaning that if you took the coordinates from a point on the tip of a guardrail from a point cloud and compared it to a survey grade GPS shot at that same point in the field it would be off +/- 3 feet. However, if you measured a point at the bottom of the guardrail and one at the top of the guardrail both in the point cloud, the distance would be very close to the distance you would find in the field. The point cloud is tight, the location is less exact. These are corrected on survey grade LiDAR by placing and surveying identifiable objects in the field and adjusting the LiDAR to these objects.

CONNECTED INFRASTRUCTURE

Big data is integrating with transportation infrastructure. In fall 2018, MnDOT deployed connected vehicle communication infrastructure in traffic signals on the Highway 55 corridor between Minneapolis and west of Interstate 494 (see figure 2).[8] The connected traffic signals sent information to

Figure 1: Trunk Highway 55 Connected Corridor between Minneapolis and I-494

vehicles with the goal of improving driving safety and efficiency.[9][10]

The information that the system broadcasts included:

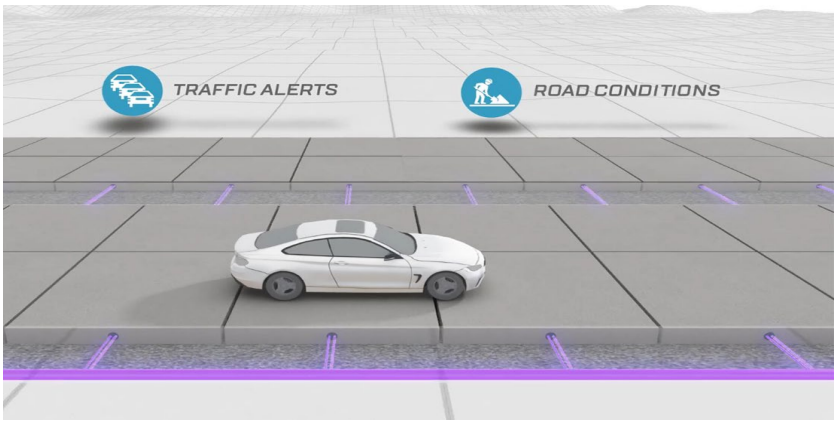
- Traffic Signal Phase and Timing: Allows vehicles to receive information about upcoming traffic lights and when they will change.

[8] Minnesota Department of Transportation, "Connected Corridors," Traffic Engineering, http://www.dot.state.mn.us/its/projects/2016-2020/connectedcorridors.html (accessed February 6, 2020).
[9] Ibid.
[10] The connected corridor technology is called Dedicated Short Range Communications (DSRC). It is similar to radio technology that is in most vehicles today. The main difference is that vehicles have to tune to the right channel and be able to listen for traffic signals sending messages. Vehicles also need to be close enough to transmitting radios to receive messages. Vehicle manufacturers are exploring ways to use 5G technology in new cars to increase the range of DSRC.

- Eco-Driving Data: Collaborates with research applications to send real-time traffic information to researcher drivers to improve fuel economy and driver awareness.
- Snowplow Signal Priority: Reduces signal stops for snowplows allowing for faster snow removal and less fuel use.

The Highway 55 connected corridor was built secure by design, rather than as an afterthought. From the start, MnDOT rolled out the technology with security by design. The corridor had a series of preventative measures, and MnDOT partnered with a third-party company to certify transmitted messages to protect against cyber threats. At the time and still today, most vehicles on the road do not have the technology necessary to use the connected corridor. This will likely change over time.

Figure 2: Integrated Roadway Visual Showing Embedded Sensors in the Road



Other states such as Florida and Wyoming are also developing connected corridors. The Colorado Department of Transportation (CDOT) embedded sensors in their roadways. CDOT partnered with Integrated Roadways to pilot" precast and interlocking concrete slabs embedded with an array of sensors, processors and antennae" (see Figure 3).[11] The smart pavement tracks the flow of vehicles and can alert emergency response personnel of vehicles that have gone off the road. The sensors collect data on vehicle position, weight, velocity and navigation.[12] This information would be shared directly with vehicles. Like Minnesota's connected corridor, the infrastructure broadcasts information to help drivers increase their fuel economy, safety and awareness.

GOVERNMENT DATA SHARING

The open data movement is encouraging the publishing of data free-of-charge to increase transparency, elevate data standards, and maximize public investment in systems that produce data. However, governments need to balance data sensitivity and sharing.

[11] Integrated Roadways, "Our Story," http://integratedroadways.com/ (accessed February 6, 2020).
[12] Ibid.

Since 2010, MnDOT has managed the open data source known as Minnesota Geospatial Commons—a "collaborative place for users and publishers of geospatial resources about Minnesota."[13] Currently, thirty-two organizations contribute data. These include the University of Minnesota, counties, state agencies, and transit authorities. Researchers, cartographers, web and application developers, planners, journalists and interested citizens use the Commons.

In some cases, private companies download free public data and repackage it to make a profit.[14] For instance, AccuWeather, a private company that produces forecasts, uses National Oceanic and Atmospheric Administration (NOAA) data and projections. There has been tension between NOAA and private weather forecasting companies over the years because of their desire to privatize NOAA's data and become the sole producer of forecasts.[15,16] As governments invest in smart city technology such as connected corridors, there may be pressure to privatize that data. This has implications for maintaining individual privacy and has the potential to hinder the open data movement.

Governments need to consider data transparency and security tradeoffs. For instance, public transit data can inform planning decisions. However, to illustrate data security concerns, fare card registration information (names, addresses, and credit card numbers) can be stolen if data is not anonymized and stored correctly.[17] In 2020, New York City subway began integrating OMNY— a tap payment system used at turnstiles via a debit card, smartphone, or reloadable card. The system collects smartphone device identifiers, location, and transaction information. It is effectively possible to track riders' daily movements down to the minute.[18] Privacy advocates around the nation are pushing for more government accountability and disclosure of the data collected.

[13] Minnesota Geospatial Commons, "About the Minnesota Geospatial Commons," https://gisdata.mn.gov/content/?q=about (accessed February 25, 2020).
[14] Anthony Williams, "Maybe Government Data Shouldn't Always Be Free," *City Lab*, April 16, 2017. https://www.citylab.com/life/2017/04/maybe-government-data-shouldnt-always-be-free/523095/ (accessed February 12, 2020).
[15] U.S. Congress. Senate. Committee on Commerce, Science, and Transportation, *Nomination of Barry Lee Myers to be under Secretary of commerce for Oceans and Atmosphere; and Administrator, National Oceanic and Atmospheric Administration (NOAA).* 115th Cong., 1st sess., 2017, 115-647. https://www.congress.gov/115/chrg/CHRG-115shrg37230/CHRG-115shrg37230.htm (accessed August 24, 2020).
[16] Jonathan Porter, *Comments of Accuweather: Allocation and Service Rules for the 1675-1680* (Washington, D.C., 2019). https://ecfsapi.fcc.gov/file/1062173865405/AccuWeather_Comments_WT_Docket19_116.pdf (accessed August 25, 2020).
[17] Cecilia Viggiano et al., *Data Sharing Guidance for Public Transit Agencies – Now and in the Future* (Transportation Research Board, 2019).
[18] Ali Winston, "The NYC Subway's New Tap-to-Pay System has a Hidden Cost — Rider Data," Verge, March 16, 2020. https://www.theverge.com/2020/3/16/21175699/mta-omny-privacy-security-smartphone-identifier-location-nyc (accessed Apri 27, 2020).

CYBER SECURITY IN TRANSPORTATION

TRADITIONAL TRANSPORTATION OPERATIONAL SYSTEMS

TRAFFIC MANAGEMENT SYSTEMS AND ROADWAYS

Transportation systems create a variety of cyber security concerns. Since 1986, the United States Department of Transportation (USDOT) has supported initiatives to get intelligent transportation systems (ITS) on roadways.[19] Signal controls, digital communications, and tolling systems are some examples (see Figure 4). The US Federal Government started exploring the implications of cyber security on roadways in 2012. The National Intelligent Transportation Systems Architecture guideline provided a framework for the consistent security-conscious deployment of ITS.[20]

Figure 3: Dynamic Message Sign on Interstate 35W South

A potential security and roadway management technology is blockchain. Blockchain was invented in 2008 to serve as a public ledger for cryptocurrency Bitcoin.[21] One of the advantages of blockchain and why tolling agencies are interested in adopting the technology for the roadway is because it decentralizes transactions between individuals and tolling agencies. When someone uses a debit card to purchase an item at the store, they

[19] Some of the policies that push for ITS are: 1986 Intelligent Vehicle Highway System Initiative (USDOT), this was amended in the 1991 ISTEA legislation as Intelligent Transportation Systems (ITS) and has since been revised and updated in every federal transportation act since. National Cooperative Highway Research Program – Transit Cooperative Research Program, *Protection of Transportation Infrastructure from Cyber Attacks: A Primer* (Transportation Research Board, 2015).

[20] Ibid.

[21] A cryptocurrency public ledger is a decentralized, record keeping system that tracks transactions between participating partners. Participants verify each transaction for authenticity. Rajat Rajbhandari, "Exploring Blockchain – Technology behind Bitcoin and Implications for Transforming Transportation," *Texas A&M Transportation Institute* (January 2018).

swipe their card at the checkout counter. Businesses that process any part of a transaction are vulnerable to hackers attempting to steal customer financial information. The advantage of blockchain is that transactions are decentralized, and vulnerable information is kept separate when paying. [22] If blockchain was used for roadway tolling, a driver would create an e-wallet and purchase toll tokens to fill it. When the driver passes through the toll booth, tokens are withdrawn. Beyond customer privacy, the US tolling industry believes they can save money. They estimate an annual savings of $300 million (nationwide) on credit card transaction fees. In the long term, the industry speculates that blockchain could allow for a nationwide tolling system where individuals use a single account to pay for their road use. [23]

Blockchain technology impacts freight connected and automated vehicles, shared ownership of rideshare vehicles and protecting the IoT. [24] These benefits do not come without disadvantages. First, decentralized control of data can make it hard to regulate and enforce. Additionally, changes to the blockchain structure can make it vulnerable to some cyberattacks. Last, cryptocurrencies are susceptible to unpredictable price swings. [25, 26]

One promising innovation from blockchain technology is the potential for distance-based user-fees (DBUF). As gas tax revenues decline and vehicles become either more fuel-efficient or electric, government agencies are looking for new infrastructure funding sources. [27] A DBUF would charge drivers based on miles driven instead of gallons of gas purchased. A concern with DBUFs is that the administrative overhead would diminish or negate benefits gained. However, blockchain could facilitate DBUF and minimize overhead. Gas tax administration typically costs one percent of revenue. In contrast, studies show that administering a DBUF would fall between five and six percent of revenue. [28] Blockchain could significantly reduce the overhead costs since the technology would automatically route revenues to appropriate coffers and verify transactions without the need for administration.

Traffic tolling also uses ITS to collect payments. Whether by toll road or toll lanes, the number of states embracing tolling technology to relieve congestion and fund transportation projects is increasing (see Figure 5). [29] To reduce overhead costs and increase security, academics and numerous tolling companies are considering blockchain, a payment system that decentralizes valuable information while streamlining payments.

---

[22] Mohamed Rahouti, Kaiqi Xiong, and Nasir Ghani, "Bitcoin Concepts Threats, and Machine-Learning Security Solutions," *Institute of Electrical and Electronics Engineers* 6 (November 2018).
[23] Rajat Rajbhandari, "Exploring Blockchain," Texas Transportation Institute, January 2018. https://rosap.ntl.bts.gov/view/dot/34863.
[24] Ibid.
[25] Mohamed Rahouti, Kaiqi Xiong, and Nasir Ghani, "Bitcoin Concepts Threats, and Machine-Learning Security Solutions," *Institute of Electrical and Electronics Engineers* 6 (November 2018).
[26] Caitlin Long, "Bitcoin, the Dollar and Facebook's Cryptocurrency: Price Volatility Versus Systemic Volatility," *Forbes*, June 29, 2019. https://www.forbes.com/sites/caitlinlong/2019/06/29/bitcoin-the-dollar-and-facebooks-cryptocurrency-price-volatility-versus-systemic-volatility/#45727fbd88b8 (accessed February 12, 2020).
[27] University of Minnesota, "Distance-Based Mileage Fees could be Tested through Shared-Mobility Providers," (Center for Transportation Studies, 2019) http://www.cts.umn.edu/publications/catalyst/2019/january/fees (accessed February 24th, 2020).
[28] Paul Sorensen, Lisa Ecola, and Martin Wachs, *Mileage-Based User Fees for Transportation Funding: a Primer for State and Local Decisionmakers* (RAND Corporation, 2012) 15.
[29] David Schaper, "More States Turning to Toll Roads to Raise Cash for Infrastructure," National Public Radio, January 18, 2018. https://www.npr.org/2018/01/18/578865204/more-states-turning-to-toll-roads-to-raise-cash-for-infrastructure (accessed February 12, 2020).

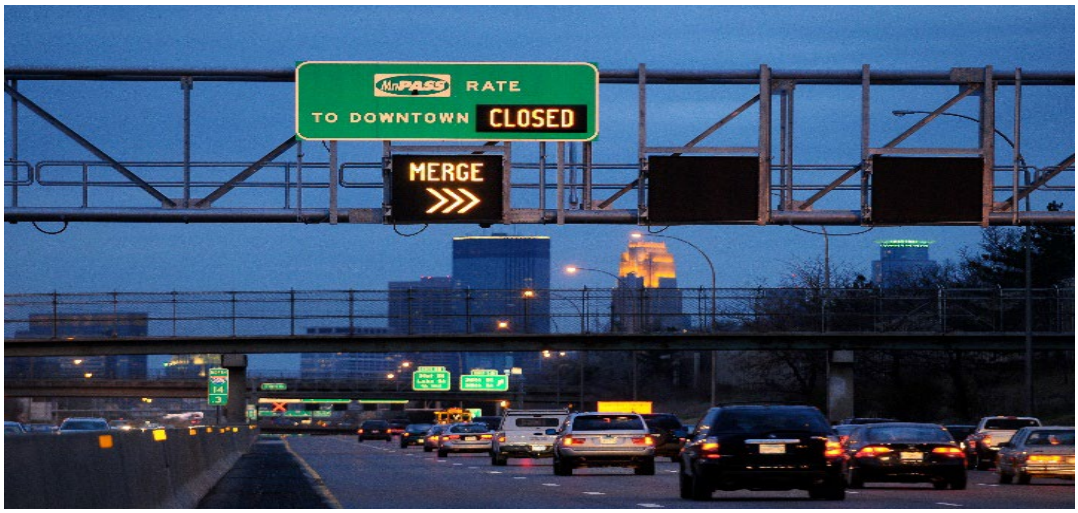


Figure 5: Managed Lane on 35W in Minneapolis

Another area of cyber security is in transportation management centers (TMCs). TMCs oversee transportation management infrastructure such as dynamic message boards and traffic signals. They also coordinate responses to roadway incidents, and work with regional partners in government and the media.[30] TMCs' risks include malware[31] attacks that tamper with signals, dynamic signs and connected infrastructure.

In the Netherlands, there is a publicly available smartphone app that cyclists can use to alert traffic signals to their approach and prompt the signal to allow them through. In 2020, researchers hacked this app to see if they could get signals to change without the presence of a cyclist. They were able to do it successfully in nearly 10 Dutch cities.[32] If the app could be used to manipulate signals outside of its original purpose, it could create widespread traffic and confusion.

PUBLIC TRANSIT

Rail systems have similar security threats as roadway networks. While new light rail systems are designed with Wi-Fi, older freight and commuter rail lines are finding ways to integrate ITS to manage speeds and traffic signals.[33] The challenge with older train networks is they were built before cyber security was a concern. There are likely less layers of protections and fail-safes if an attacker breached a new wireless network used with an old train network. Modern transit systems have layered protection against hackers. For example, modern light rail uses

[30] Ibid.
[31] Malware is "a software or firmware intended to perform an unauthorized process that will have adverse impact on the confidentiality, integrity, or availability of an information system. A virus, worm, Trojan horse, or other code-based entity that infects a host. Spyware and some forms of adware are also examples of malicious code."
National Institute of Standards and Technology, "Malware," United States Department of Commerce, https://csrc.nist.gov/glossary/term/malware (accessed April 6, 2020).
[32] Andy Greenberg, "Dutch Hackers Found a Simple Way to Mess With Traffic Lights," *Wired*, August 5, 2020. https://www.wired.com/story/hacking-traffic-lights-netherlands/#:~:text=Their%20hack%20would%20spoof%20nonexistent%2ccross%20in%20a%20perpendicular%20direction. (accessed August 25, 2020).
[33] Rosie Perper, "Railway Systems Could Be Hackers' Next Big Target – and Derailing Trains Wouldn't be that Hard," *Business Insider*, May 18, 2018. https://www.businessinsider.com/cyber-attacks-targeting-railway-systems-next-2018-5 (accessed February 12, 2020).

Automated Train Protection monitoring. The system scans for potential issues and can instantly stop a compromised train.[34]

Public transit agencies are also adding benefits such as free onboard Wi-Fi and real-time transit tracking to attract new riders. These new systems create cyber security risks. If these tools are not adopted with a commitment to cyber security, hackers can enter the network, disrupt the train, system, and even steal employee data and customer transit card registration.[35] Any part of the system that uses Wi-Fi is theoretically vulnerable.

CONNECTED AND AUTOMATED VEHICLES

The emergence of connected and automated vehicles raises concerns for hacking and remote take overs.[36] What once might have seemed like distant future, vehicle hacking is a reality now.[37]

In the past, hackers took control of vehicle systems through hard-wired connections to computers in an automobile. With the emergence of telematics systems (internet-connected automobile services) in cars, hackers can locate, track, and connect to a vehicle remotely.[38] Common examples of telematics systems include Chrysler's Uconnect, GM's OnStar, Toyota's Safety Connect, and others. Moreover, vehicle manufacturers are progressing towards only producing electric vehicles—meaning that whenever they are charging, they are connected to the grid and vulnerable to hacking. The potential security implications are vast.

The State of Minnesota adopted its first Connected and Automated Vehicle Strategic Plan in July 2019 to answer some of these questions at the state level.[39] The plan considers the states' data management practices and assesses cyber security risks.

GOVERNMENT SECURITY AND RANSOMWARE

Governments are experiencing and paying for a growing number of ransomware attacks. In pursuit of collecting big data and utilizing smart city technology, governments are likely under-investing in cyber security. Staying up to date is expensive. To begin with, many governments often have limited staff capacity and resources. It is also politically challenging to lobby for cyber security money when attacks are seemingly infrequent. This perception is inaccurate.

In 2021, governments worldwide saw an 1,885% increase in ransomware attacks[40]. Table 1 shows the number of ransomware victims and estimated losses from 2015-2019. While this data is for all attacks reported to the FBI's

---

[34] National Cooperative Highway Research Program – Transit Cooperative Research Program, *Protection of Transportation Infrastructure from Cyber Attacks: A Primer* (Transportation Research Board, 2015).
[35] Lurae Stewart et al., *Cybersecurity Considerations for Public Transit* (Washington, DC: American Public Transportation Association, 2014).
[36] James Anderson et al., *Autonomous Vehicle Technology: How to Best Realize Its Social Benefits* (Santa Monica, CA: RAND Corporation, 2014). https://www.rand.org/pubs/research_briefs/RB9755.html (accessed February 12, 2020).
[37] Andy Greenberg, "Hackers Remotely Kill a Jeep on the Highway—With Me in it," *Wired*, July 21, 2015. https://www.wired.com/2015/07/hackers-remotely-kill-jeep-highway/ (accessed February 12, 2020).
[38] Ibid.
[39] Minnesota Department of Transportation, *Connected and Automated Vehicle Strategic Plan*, 2019. http://www.dot.state.mn.us/automated/docs/cav-stategic-plan.pdf (accessed April 6, 2020).
40 Amiah Taylor, "There's a huge surge in hackers holding data for ransom, and experts want everyone to take these steps," *Forbes*, February 17, 2022, https://fortune.com/2022/02/17/ransomware-attacks-surge-2021-report/, (accessed May 22, 2022).

Internet Crime Complaint Center and not attacks on governments, it helps to demonstrate the increasing frequency and severity of ransomware attacks. These attacks hit supply chains and can cause widespread system downtime, economic losses and damages to reputations. Understandably, the threat of ransomware attacks are an increasing concern for government agencies and the transportation industry.

Table 1: Number of Victims Reporting Ransomware Attacks and their Losses[41]

Date	Victims	Victim Losses	Average Loss per Victim
2019	2,047	$8,965,847	$4,380
2018	1,493	$3,621,857	$2,426
2017	1,783	$2,344,365	$1,315
2016	2,673	$2,431,261	$910
2015	2,453	$1,620,814	$661

While the FBI data does not include government losses, there are recent examples that can demonstrate the impacts to an agency and the public. In February 2018, Colorado DOT experienced a ransomware attack that compromised over half of their computers. It took four weeks to regain control of the network and contain the attack. While no ransom money was paid, CDOT spent $1.7 million in overtime pay and unanticipated expenses.[42] A primary concern of these attacks includes the corruption of institutional data on compromised computers—another loss.

Similar attacks occurred in Atlanta in 2018 and Baltimore in 2019. In both instances, the cities repelled the attacks but spent millions on contractors and unexpected expenses. The Atlanta breach was so widespread that sensitive data including "warrant issuances, water service requests, new inmate processing, court fee payments and other online bill-pay programs across a range of city departments" were affected.[43] The Atlanta breach even forced officials to shut down Wi-Fi at Hartsfield-Jackson Atlanta International Airport, the busiest airport in the world.

While even the most protected computer can be hacked, governments can minimize their risk. Some of the actions government agencies can do to mitigate risk include frequently back-up network computers, storing the

[41] Tamara Chuang, "How SamSam Ransomware Took Down CDOT and How the State Fought Back – Twice," *Colorado Sun*, February 2, 2020. https://coloradosun.com/2020/02/03/how-samsam-ransomware-took-down-cdot-and-how-the-state-fought-back-twice/ (accessed February 12, 2020).
[42] Ibid.
[43] Laila Kearney, "Atlanta Ransomware Attack Throws City Services into Disarray," *Reuters*, March 23, 2018. https://www.reuters.com/article/usa-georgia-cyber-idUSL1N1R51V9 (accessed February 12, 2020).

back-up on an external hard drive or separate network, and training employees on phishing[44] and cyber security threats. These actions can dramatically increase network safety.[45]

CONSUMER PROTECTION AND REGULATION

To combat increasing cyber security concerns and data breaches, a variety of actions are taking place across sectors. An agreement between 18 automakers in January 2016 determined a set of safety principles to standardize cyber security. The deal included using data to spot cyber security breaches and formed a task force to examine the cooperation model between airlines and governments—a potential example for car manufactures who might want to cooperate with governments.[46]

In 2016, USDOT outlined a series of safety principles to help automated vehicle manufacturers test and improve vehicle safety. These principles included early reporting of potential risks, vehicle recalls, and a commitment to work collaboratively across industries. Specifically, in terms of cyber security, USDOT urged automakers to focus "on layered solutions to ensure vehicle systems are designed to take appropriate and safe actions, even when an attack is successful."[47]

When it comes to consumer data protection, the State of California leads the way. As of January 1, 2020, the California Consumer Privacy Act gives individuals a series of rights "relating to the access to, deletion of and sharing of personal information that is collected by businesses."[48] Specifically, individuals have the right to know what personal information is being collected, delete personal information, opt-out of having their data sold, and the right to non-discrimination in price or services if they choose to withhold their information.[49] This policy redefines how individuals can protect their data and has implications for big data and smart city applications.

The use of big data is not free from bias.[50] Smartphone ownership can result in sampling-related biases that make mobile phone data unrepresentative of communities served by transportation planning and decision-making. Social media platforms have varying levels of use across ages, races, genders, etc. Much of the transportation data available comes through the use of surface transportation and thereby favors personal vehicles. As most

[44] Phishing is "a type of online scam that targets consumers by sending them an e-mail that appears to be from a well-known source – an internet service provider, a bank, or a mortgage company, for example. It asks the consumer to provide personal identifying information. Then a scammer uses the information to open new accounts, or invade the consumer's existing accounts." United States Federal Trade Commission, "Phishing Scams," Under Identity Theft, https://www.ftc.gov/news-events/media-resources/identity-theft-and-data-security/phishing-scams (accessed April 6, 2020).
[45] United States Department of Homeland Security, "Security Tips (ST19-001): Protecting against Ransomware," National Cyber Awareness System, https://www.us-cert.gov/ncas/tips/ST19-001 (accessed February 24th, 2020).
[46] Bernie Woodall and David Shepardson, "Major Automakers Announce Voluntary U.S. Effort to Boost Safety," *Reuters*, January 15, 2016. https://www.reuters.com/article/autos-detroit-safety/major-automakers-announce-voluntary-u-s-effort-to-boost-safety-idUSL2N14Z1K3 (accessed February 13, 2020).
[47] United States Department of Transportation, "Proactive Safety Principles 2016," NHTSA, https://www.transportation.gov/briefing-room/proactive-safety-principles-2016 (accessed February 13, 2020).
[48] State of California Department of Justice, "California Consumer Privacy Act," under Privacy, https://oag.ca.gov/privacy/ccpa (accessed February 13, 2020).
[49] Ibid.
[50] Virginia Tech Transportation Institute, "Sources and Mitigation of Bias in Big Data for Transportation Safety", https://vtechworks.lib.vt.edu/bitstream/handle/10919/88893/02-026_Final%20Research%20Report_Final.pdf?sequence=1 (accessed July 18, 2020).

transit data comes from automatic counter systems and smartcards, these data sources are inherently less biased compared to other sources such as smartphones that fail to accurately represent the entire population.

FUTURE OF BIG DATA AND CYBER SECURITY

Big data and cyber security are linked. The number of vulnerabilities are increasing every year, and at an even faster rate for those related to the Internet of Things.[51] For every IoT application that improves your driving efficiency or maps your trip home, there are cyber security and application concerns. For MnDOT, the rapid growth of third-party, open platform apps is a concern. Smartphones are not necessarily secure and can provide a backdoor for cyber attackers to access connected corridor and other transportation systems.[52]

The increasing use of big data will bring more legal and ethical concerns and potential implications for the public. The COVID-19 pandemic highlights the changing context around the use of personal data. The pandemic compelled South Korea, Israel, Iran, and Singapore, among others, to collect anonymized cellphone location data to track the spread of the virus. In particular, the Singapore app used Bluetooth technology to determine if someone was within two meters of an infected individual. The app kept a record of Bluetooth interactions, which could be used for contact tracing. Big data tools such as the app used in Singapore helped with the government response to the pandemic. However, tracking individuals' real-time data locations is a significant risk to personal privacy.

In addition to biases in data collection, the use of data also presents potential equity concerns. A study completed by National Institute of Standards and Technology found that the ability to match two images of the same person varies by demographics occurred in a majority of algorithms.[53] While the cause is unknown, the study rose concerns about the limitations and appropriate use of such algorithms. Instances of "false positives"—when a face recognition system wrongly considered photos of two individuals to be the same person—is higher for women, older adults, children, and individuals with darker skin tones.[54] As the volume and application of big data grows, so do protection and regulation concerns for those individuals whose data has been collected.

Cyber security and big data are intertwined. While big data presents a variety of advantages and opportunities, insufficient data management, varying levels of accuracy, biases, and security protocols make governments and individuals vulnerable. Smart policies that systematically consider implications can maximize big data advantages and minimize concerns.

RELATED TOPICS

- Automated Vehicles
- Electrification & Alternative Fuels

[51] IBM, X-Force Threat Intelligence Index 2022, https://www.ibm.com/downloads/cas/ADLMYLAZ. (accessed May 22, 2022).
[52] Andy Greenberg, "Dutch Hackers Found," *Wired*, August 5, 2020.
[53] National Institute of Standards and Technology, "NIST Study Evaluates Effects of Race, Age, Sex on Face Recognition Software," December 19, 2019, https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software, (accessed May 22, 2022).
[54] National Institute of Standards and Technology, " Facial Recognition Vendor Test," https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf (accessed July 18, 2020).

- Mobile Technology
- Shared Mobility

Minnesota's vision for transportation is known as Minnesota GO. The aim is that the multimodal transportation system maximizes the health of people, the environment and our economy. A transportation vision for generations, Minnesota GO guides a comprehensive planning effort for all people using the transportation system and for all modes of travel. Learn more at MinnesotaGO.org.

REVISION HISTORY

Date	Summary of revisions
May 2016	Content previously included in the Sensors, Monitors & Big Data paper that has been archived.
August 2022	Updated with new data and information.