CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES

TECHNOLOGY POLICY PROGRAM

April 2017

IoT, Automation, Autonomy, and Megacities in 2025: A Dark Preview

Michael Assante and Andrew Bochman[1]

Preface

This paper extrapolates from present trends to describe plausible future crises playing out in multiple global cities within 10 years. While predicting the future is fraught with uncertainty, much of what occurs in the scenarios that follow is fully possible today and absent a significant course change, probable in the timeframe discussed.

It is not hard to find tech evangelists touting that ubiquitous and highly interconnected digital technology will bring great advances in productivity and efficiency, as well as new capabilities we cannot foresee. This paper attempts to reveal what is possible when these technologies are applied to critical infrastructure applications en masse without adequate security in densely populated cities of the near future that are less resilient than other environments. Megacities need and will deploy these new technologies to keep up with insatiable demand for energy, communications, transportation, and other services, but it is important to recognize that they are also made more vulnerable by following this path.[2]

To illustrate what these eventualities could look like, we have constructed four scenarios for the not-too-distant future (2025) that lay out some of the more extreme risks we may face in an all-digital world.

[1] Authors' caveat: we are not commenting on specific organizations or technology deployments.

[2] The authors are fans of predictive technologies and machine learning systems—given growing populations and scarce resources, these technologies will be essential for our collective future. This paper is meant to make people aware that even well-balanced systems will be susceptible to perturbations that can grow large as we centralize learning and connect the many things. System owners, city planners, and technology communities and innovators need to consider these in their designs and investments. We also want the world's innovators to take notice of the opportunity to change these stories. They can be different if defenders can harness the power of the analysis to quickly determine unauthorized changes (tampering and experimentation) and use machine learning to spot patterns. Our call to action is for innovators to make corresponding investments and advancements to leverage the breakthroughs in cognitive systems and data and apply these capabilities to analyze and predict security-related outcomes.

Introduction: Setting the Stage

Looking back, we can discern the root causes of the events we're about to summarize. There was lots of excitement about what the Internet of Things (IoT), as it was called back then, was going to do for humanity. Here's Marc Andreessen, one of the first Internet pioneers, thinking out loud and with no small amount of optimism, in 2016:

> The photos are all going, "Hey," and the plate goes and refills itself and brings you fresh food, and your beer mug tells you you're drinking too much. Everything is just smart. This is my view of the Internet of Things: you're able to infuse intelligence into everything, you're able to put a chip in everything, you're able to put software in everything, you're able to connect everything online and just everything is a lot smarter. The doorknob is a lot smarter, and the lightbulb is a lot smarter, and your wristwatch is a lot smarter. Everything starts to get really, really smart.

On the role of technology in improving the human condition, techno utopianism has been soundly besting ludditism going on two centuries now, and the world of late 2025, with its autonomous vehicles, fully integrated smart cities, deep virtual and augmented realities, and artificial intelligence getting ever closer to human-parity general intelligence, is the result.

With that said, what's transpired around the world just this past year has got to give pause to even the most ardent optimists. Compiled below are the unnerving events witnessed in four of the planet's largest and most important cities as reported by the local media and then deconstructed by an artificial intelligence (AI)-assisted omniscient cyber forensicist. In all it's clear that the technologies that helped the infrastructure managers of these cities handle the almost incomprehensibly complex operations of a modern megacity were also the root cause (or at a minimum, the enabler) of the disasters that befell them. Undoubtedly, cyber attackers played a greater or lesser role in getting these crises rolling, but it appears that despite decades of warnings, in the name of progress we've made things ever so easy for them. Now all we can hope for is that we learn from these experiences and implement changes as quickly as possible, knowing full well that change in infrastructure matters never comes quickly.

City Scenario I: Bangkok

Dispatch

Bangkok Post, Wednesday, April 23, 2025, Midnight—On a normal day, most residents of Bangkok could expect clean water to flow from their faucets and their toilets to flush. Bold infrastructure engineering work done in the late nineteenth century supported public health improvements that led to ever-larger populations. Population growth in turn put stress on the systems and was relieved by subsequent waves of engineering imagination and excellence. Some of the world's largest, most efficient water treatment plants have given this city some of the most affordable, mainly clean water in Asia.

1616 RHODE ISLAND AVENUE NW
WASHINGTON, DC 20036
TEL. (202) 887.0200
FAX (202) 775.3199

However, as of five days ago, nothing in this sprawling city of 30 million has been normal:

- Multiple fires raging out of control when fire hydrants couldn't produce water
- Industrial businesses shuttered because they couldn't make their products without reliable water supplies
- Foul water coming out of faucets, spilling forth from toilets in apartments and on the streets from manhole covers
- And perhaps worst is that some power plants in and around the city are running at reduced capacity due to having less of the water they require for cooling, and power outages are undermining all efforts to restore order.

After years of droughts brought groundwater to historically low levels, a few days ago reports of low and then no water pressure started coming in from households, businesses, and government offices. By early evening the state-run Metropolitan Waterworks Authority (MWA) issued a statement saying it had lost control over the majority of the pumps responsible for maintaining water pressure, as well as its operator consoles, and was investigating the issue.

Throughout this ordeal, the governor of Bangkok tried to keep a calm face. But today he appeared to lose his courage. On the Royal Thai Army's Channel 7, the mayor said:

> People of Krung Thep, I realize many of you cannot hear my message, but for those who can, I strongly urge you to be strong, and carry on with as much faith and discipline as you can muster. It appears we may be the victims of an unprecedented cyber attack on our water infrastructure. The smartest engineers in our city are working day and night to understand the full extent of the attack and with luck, will restore water, electricity, order, and hope to our city as soon as possible. Otherwise, I am not sure what will become of us.

Tomorrow will be day six. With order breaking down, the Army trying to help the police keep the growing riots in check, and with bottled water reserves almost depleted, we can only pray the engineers will have success soon.

Omniscient Forensic Analysis

The roots of the problem began earlier that April when a water authority engineer received a file from a collaboration platform used for getting new files from the city's primary water service infrastructure automation vendor. The file appeared to correspond to a firmware update posted on the vendor's knowledge portal indicating it was required to patch a memory leak problem. Once downloaded, the file was then transferred from the engineer's laptop to three engineering workstations.

It only took seconds for the attackers' command-and-control system to find the signal emanating from the Trojan contained in the file. Almost immediately, additional implants made their way

undetected on to the target workstations and began to exfiltrate the information needed to seed the necessary changes to system software. That was the software residing on actuators and digitally controlled pumps throughout the sprawling water system serving central Bangkok and surrounding districts.

While remaining undetected, the attackers eventually learned enough to capture the digital credentials they needed to manage the IT and operational technology (OT) infrastructure. Data stolen from both the business network and the water Supervisory Control and Data Acquisition (SCADA) network provided the keys needed to focus the attackers' engineering efforts. It took three months, but testing proved their bricking[3] payloads would load successfully 90 percent of the time and result in irrecoverable device shutdowns. Staging the automatic software loads was now the only step to be completed.

When the first attack came, the Bangkok water system experienced several waves of destruction as malicious firmware was propagated to digital systems, including variable-speed drives required for pumps and communication devices throughout their water transmission and distributed system. Engineering teams were getting good at using analytics to predict failures and deal with probabilistic failures in the system, but the scale of these failures was never seen before. Equipment failures spanned from distribution pumping stations, water treatment plants and chemical feeding systems, to transmission pumping stations. The attackers were able to shut down pumping at five key stations rapidly depressurizing the entire water distribution system and setting off an overwhelming onslaught of alarms at the water control center. Programmable logic controllers (PLCs) began to report errors before their symbols went gray on operator consoles. First the pumps, then the routers and modems, and finally the controllers were lost.

Work crews were unable to quickly repower units, to say nothing of the systems that were in a bricked state. System planners ran through dusty plans for restoring the system by using older pumps found in outlying stations. A crisis soon developed as newer systems to measure water quality were no longer feeding data up to the quality analysis application running in the private cloud. The water-quality checkpoints failed to report data and vital components failed in the elaborate array of automatic chlorine feeding systems. The overflow of sewage was now threatening water quality throughout the system.

Not only were they blinded, but the operators were robbed of the tools necessary to control water processing for treatment and pumping. With PLCs no longer functioning, there were also problems with digital actuators such as discharge and suction valves, variable-speed drives, motor-control units, and supply and exhaust fans. This was an attack on a previously unprecedented scale, and the IT group and SCADA support engineers found they did not have the tools for the job or the ability to touch the staggering number of affected devices. The small IT department was unable to

[3] "Bricking" is a term used when a computational device is rendered inoperable or is unable to perform its intended function. A bricked device would be described as entering a disabled state. A "disabled state" encapsulates any behavior that deviates from the documented function of the device. Examples can include nonresponsive connectivity ports, improper I/O function, erroneous status information, or communication by the device that it is in a faulted state.

keep up with the reporting and requests for assistance from the Raw Water Development Department, Treatment Plant Services, and Water Distribution and Control Departments.

Frantic calls to device manufacturers became an all-hands effort as inventories were quickly exhausted. Devices on the shelf had already been rushed to a central station nearest to the city's emergency response center and sports arena.

This cyber-induced crisis had the following roots:

1. Insecure remote connections facilitated the attack
2. Inability to detect intrusions allowed attackers to discover many firmware devices and to engineer payloads for several different models, allowing for a massive attack
3. Automation and connectivity provided a pathway to find, touch, and deliver firmware uploads
4. Firmware loading lacked advanced authentication mechanisms

City Scenario II: Shanghai

Dispatch

Xinhua News Service, May 5, 2025, 11:10 pm—At the time of this report, 80 percent of Shanghai's transportation system is completely inoperable. The computer systems that manage airports, airlines, trains, subways, buses, and more have been massively disrupted. Airlines are reporting their logistics scheduling systems are unstable. The few rail operators we reached are saying they can't see the positions of their trains and, in some cases, can't verify the position of their track switches. Following established procedures in this state of uncertainty, they stopped all movement. To top it off, bus and taxi services, both autonomous and with drivers, are unable to keep up with the unprecedented surge in demand and may be experiencing glitches of their own.

Although the cause is unknown, some opinions are forming. According to Mr. Steve Hu, chairman of Huawei's Global Cyber Security and User Privacy Committee:

> The scale of this attack on transportation infrastructure seems unprecedented. There can be little doubt there is a nation state behind this action. Who else could muster the resources to create so many concurrent impacts on such diverse systems?

In time, we'll come to know how fast these services can be returned to normal and hopefully identify the root cause. What we do know for sure: hotels are reporting they are completely full, more than 3 million people are stranded, and it's going to be a long night.

Here's a recap of today's events. During this evening's rush hour, subways and trains serving Shanghai's Pudong and Hongqiao international airports started running late and then stopped

running altogether. In short order, concentric rings of similar troubles spread across the greater Shanghai region. Rail commuters, both residents of the city as well as business people and tourists from other parts of China and around the world, are utterly stranded. Buses and taxis initially responded to the huge surge in demand; however, as the 15 million people dependent on trains and subways turned to these alternatives, they too experienced systems failures that rendered them nearly useless. One international banker we interviewed said he's never seen anything like this:

> I was waiting for the 4:15 train to Pudong and even though the monitor said it was arriving, it never actually came. Eventually I tried hailing a taxi but the app indicated a five-hour wait time. My flight to Chengdu for work was supposed to depart at 7 pm but now I understand it was just canceled. I give up. I want to just go home but even that now seems impossible.

Then came the airlines. Chinese airlines including Air China, Shanghai Airlines, and Juneyao Airlines as well as foreign carriers Delta, Emirates, Singapore Airlines, and others had in recent months begun reporting intermittent issues with their scheduling and logistics systems. As late afternoon turned into evening, a clear disruption to air operations led some experts to suspect a coordinated cyber attack.

The global economy took notice with sharp drops in the Shanghai and Hong Kong indexes and overseas the DOW and FTSE are falling as well in pre-opening trading. The costs to productivity seem likely to be massive, as are the ripple effects of unprecedented supply chain disruption.

Omniscient Forensic Analysis

Though contemporaries said the cause was "unknown" but attributed it to terrorist attack, in retrospect the cause was obvious and not malevolent. High-precision time measurement matters more than ever in 2025 as larger, more interconnected systems rely on the efficient exchange of accurate time-stamped data. Many developers had been warned to select their algorithms and libraries carefully, but not all heeded that advice, and this cascading transportation disruption began in 100-nanosecond increments before it built into a time typhoon:

- Dynamic power management (DPM) had been rolled out to reduce the costs of paying for the "electrification of everything"
- Shanghai had witnessed power consumption increasingly shifting from households to collections of more and more things.
- Widely deployed DPM schemas were used to shut down devices when they were not needed and to wake them before they were needed to receive/send data or process data
- Industrial Internet of Things (IIoT) implementations had been coded to optimize the performance of associated devices in an attempt to manage out inefficiencies. A software update addressed a few known bugs and added an innovative new way to manage DPM

- Recent modernization projects allowed the world's most-used metro to squeeze additional capacity from the fixed core system and already maxed-out train car-per-track arrangement

- New optimization software had taken advantage of cheap slap-on instruments to measure activity in stations and along tracks to handle growing commuter numbers

The inflows from Maglev stations and hand-off stations to other forms of transportation like airports were synchronized to better control train traffic. The software had already provided results and slight tweaks were showing more promise under incredible demands to do more with what was in place. Additional software-based controls allowed system designers to deal with the scale of more data inputs and larger sensor deployments. "Run trains closer together, safely" was the motto and driving force for innovation to include upgrading track-positioning sensors, from passive radio frequency identification (RFID) tags to more powerful multisensor devices that could include measurements that not only conveyed location, but indications of train loading and maintenance information. The software was used to coordinate messaging and device power-on based on advances in predictive analysis and being able to estimate train location while using the sensors to verify and report. All of this meant Shanghai could keep up with its growing population and continue to serve as an engine for growth and global investment.

The first failures in trackside instruments, caused by a compounding error that began impacting instruments weeks after the update had been loaded, were being handled by logic in trackside controllers and local system estimators. The predictive algorithm worked well, but its insatiable appetite for data would finally go unmet as trackside devices failed to wake in time to provide anticipated reports. Trackside controllers could not send necessary outputs and the matching of train controller-fed locational data began to deviate. The complexity of multiple data sources and the management of large underground deployments of firmware-based devices had been moved into software. The DPM software tweak left devices in a sleep state too long resulting in unanticipated extra controller-initiated communications when devices awoke outside of their predictive windows.

The predictive applications began to fail and safety logic brought trains to a stop until sufficient data would allow for the verification logic to solve. The loss of vital data and disruption in train service quickly cascaded to other transportation elements as passengers became stranded, stations were occupied to capacity, and data flows between the transit systems and other systems warned that something was terribly wrong. The larger transportation system-of-systems began to fail as humans were not where they were supposed to be and data triggered verification routines. Tremendous amounts of data being sent by automated systems and individual customer requests for automobile sharing services were overwhelming dispatching applications causing a denial of service and timely processing.

Human override of the train safety logic in the applications was ruled out and initial forensics was able to uncover the power management issue. A fallback version of the software was staged and deployed, but the scale of the instrument failure meant hours were going to spread into tens of

hours and possibly multiple days. Transit authority maintenance crews had never had to touch so many devices that quickly. Offers to have military units available to aid in the loading of fallback software were turned down as the loads were tricky and had to be verified.

If the rippling impacts of stranded commuters were not enough, the congestion of the city's cellular network began to stress priority service schemes, eventually leading to network latencies as voice data and digital messaging began to overwhelm towers and backbones. The technology implemented to digitize infrastructures had outpaced the cell networks they relied upon. The traffic models had not anticipated a day anything like this and although the cellular network remained available the latencies affected smart grid meters and telemetry signals from field terminal units and digital sensing devices. The congestion resulted in local power management conflicts that resulted in losing power to sections of the distribution system feeding one of the airports and associated operational data centers.

The loss of power prompted power meters and non-power IIoT/IoT devices to send "last will and testament" messages using capacitors for a last joule of energy. The resulting communication surges piled onto the already-congested network. More congestion resulted in additional spot power outages. The power disruptions were exacerbated by failures in backup generator and micro-grid supplies, also due to cellular network congestion. The dependencies between applications, data, and infrastructures became painfully obvious. It may be many weeks, if not months, from now before the true chain of events can be mapped out.

City Scenario III: Mexico City

Dispatch

Radio Fórmula Cadena Nacional, July 26, 2025—In other big cities around the world we've seen cyber-attacks on infrastructure spark the devolution of city services and, almost instantaneously, civil order. What's playing out here in Mexico's beloved capital is a reversal of that sequence, with all-too-familiar city employee strikes that have slowed the city to a crawl for the past few months setting the stage for something quite out of the ordinary. We Mexicans long ago learned to expect and tolerate near-crippling bureaucracy and inefficiency. But Mexico City, also known as Ciudad de la Esperanza, or "The City of Hope," has been the exception in many ways: exuberant, business-fueled perpetual motion despite the enervating friction of its incorrigible, corrupt, and bloated government.

All that, however, seems to have unraveled quickly when tens of thousands of strikers and other protesters, enraged by the latest round of pay cuts, turned out on the streets and brought the city to a weeklong standstill. Incessant social media campaigns in support of the strikers were to be expected as were cyber-attacks of mixed success on city government websites. But when the built infrastructure started acting as if it were possessed by demons, it became clear that a more disruptive type of cyber assault might be occurring. It appears now that over the course of approximately 90 minutes, about half of all the elevators in the city froze, often stuck in between floors, stranding hundreds, maybe thousands, of people all over the city in truly desperate

situations. How the cyber protesters were able to make this happen is anyone's guess. One thing is certain though: first responders including police, fire, and assorted facilities engineers were 100 percent occupied when the next crisis hit.

Within a few hours of the start of the elevator troubles, fire alarms and sprinkler systems activated inexplicably in other buildings, forcing their occupants into the street. One could sense the beginning of mass panic. With the police fully engaged in frantic rescue attempts across the city and the military not yet activated, the streets began to boil. It was at about this time that the attack on Santa Úrsula's Estadio Azteca turned out the lights, emptying tens of thousands onto already-jammed streets. The stampedes and barbarism that ensued and expanded from there have left many thinking there may be no imaginable point of return.

Omniscient Forensic Analysis

Over the last 15 years the operations and maintenance of heavily used machines like elevators and escalators have been brought into cloud analytic platforms with remote access, diagnostics, and predictive maintenance. Elevators and escalators are typically out of service two days per year as a result of planned inspection and maintenance or a malfunction. The collection of diagnostic data combined with predictive analytics and remote access provides more efficiency in servicing and enhancing the already high levels of availability. Instruments collect data and feed it over wireless pathways to communication gateway devices to then reach a controller and head up to a cloud platform. The cloud platform software provides a view to remotely manage hundreds of thousands of machines while collecting data from millions of sensors. The local building managers are able to receive a feed of the transport systems in their facilities and service providers and manufacturers can monitor machine health and maintenance for an entire fleet. The software aids engineers in determining if and when technicians need to be sent out, while equipping them with information for tests and the work that needs to be performed. Sensors can provide vibration, speed, and temperature data to building managers and service technicians' smart phones and tablets armed with maintenance applications.

These global systems harness powerful core analytics, saving money by improving computing efficiency and machine performance. Engineers on different continents can diagnose faults and performance irregularities on large numbers of machines each. Centralization has increased productivity of service providers, but it also provides an adept individual or group with the ability to remotely interact with many machines at once.

A group of hackers began toying around and found easy ways to make money using their skills. They were smart and stayed below the radar for the most part. The group acted more like a club than a gang. The recent social tensions had been a big topic at the lot gatherings. Two of their members had been experimenting with their apartment building's automated systems. They found it comical that wireless network broadcast would advertise central elevator data and west side cargo elevator. Their explorations brought them into contact with sensor data streams and a host of IP addressable microcomputers. Some had web interfaces others did not. It was mostly just for

fun until their explorations uncovered remote connections, and evidence of interactions that came from mobile phone applications and a central data depository.

When three of the club members were caught up in the strike, they began encouraging other members to get involved. One of the members was put into the hospital at the hands of riot control police, and finally the group came together to plan an assault. They used their own apartment building access to figure out how to access the systems of buildings surrounding the protests. The idea was simple: dump more people into the streets so the police would need to pull back and city officials would be forced to negotiate with the strikers.

This handful of hackers fully appreciates the potential consequences of their actions. Their actual plan was basic: put elevators into shutdown and maintenance modes while removing or changing IP addresses and configurations so machines could only be restarted onsite. The only thing holding them back was scale. They had caught the user name and password for three buildings but some implementations had not been accessed recently. A Google search yielded a hardcoded user account made by the manufacturer. Then they were in all over the city. Soon, they were knocking machines off line by the hundreds. The next move was a little more interactive as the group tripped evacuation alarms from a cracked building management application. The alarms got people moving, but it was the triggering of the fire-suppression systems based on bad data inputs to temp sensors that finished the job.

A few hours of play created this chaos and it proved to be enough to tip the city into a prolonged and brutish emergency with deadly results.

City Scenario IV: New York City (the Grand Finale)

Dispatch

New York Times, Friday, September 19, 2025, 2 pm—In what is already viewed as the worst attack on New York since 2001, and what may turn out to be many times worse before it's over, the city has just been hit with what appears to be a coordinated cyber-physical attack of the kind national security experts have been warning about for decades. The ultimate costs and causes may never be known, and it seems the largest and most famous American city will never be the same.

The city of 25 million has just been plunged into what is inarguably its worst blackout of all time. Historical blackouts (most famously in 1965, 1977, and 2003) were contained to between one and several days in duration. The current blackout is going into its third full week. All five boroughs are affected, along with parts of New Jersey, White Plains, and Long Island.

Electricity outages quickly impaired other critical city services like water and sewage, transportation, communications, and more. Residents with the means to have been streaming out of the city since July 5 and flooding suburbs to the north and west, as well as inundating Boston, Baltimore, and Washington, D.C. Drone footage captured this morning offered views that were nothing short of apocalyptic: stores shuttered and/or looted, street lights out, nonfunctional

subways, and gas cars moving fast to avoid organized bands of thieves. New York City, traumatized once more, is now held together by the National Guard acting under State of Emergency authority.

The day prior looked like it was going to be a typical Friday leading into an Indian Summer weekend, albeit a hotter one than normal, as summer high temps had been well into the 100s. Then from most accounts, the 4G and 5G phone and data networks stopped cold, and not just for residents, but for most businesses and government workers too, and the city shifted with startling speed from a festive holiday mood to anxiety and then what lies beyond anxiety.

One might have thought there'd be strong backup systems for the wireless systems on which so much depended, but one public department of public services (DPS) employee we reached shared:

> If your backup plan for loss of cellular communications is different cellular communications, then you have no backup plan . . . and that's been the plan for years now.

Others put blame on the less-than-reliable renewable energy systems that have been deployed en masse since the NY REV grid modernization plan took full effect in the late teens and early twenties. But then the hydro power the city has relied on from Canada should have saved the day, right? Well it most certainly has not.

It's hard to say where this is going to end up. The U.S. economy is in shock and the DOW and global stock markets have declined between 30 and 50 percent since July 7. Right now, your best bet for New York is to get out if you're there and stay out if you're not. The city that never sleeps has been plunged into a deep coma from which it seems unlikely to ever fully reemerge.

Omniscient Forensic Analysis

The engineers could not fathom how they ended up here. The first quarter of the twenty-first century ushered in the smart grid, which soon gave rise to the industrial internet and distributed intelligent microgrids, all of which were coordinated by the cloud-hosted computing grid. The tremendous complexity of this system of systems was concealed by mathematical algorithms and data analysis. Beautiful pictures and data displays would tell us where to go and what to do to maintain it. Modern society became tethered to a digital infrastructure that could not be catalogued. This digital fabric spread across the globe and was deeply embedded in all things, from the removal of waste water to the determination of how billions of people might best travel to work each morning. Cloud computing further united parts of the world. Many argued that globalization in the physical world had taken a step back in the 2020s, but the cyber world told a different story.

Some experts had warned about the uniquely potent risk posed by highly targeted and sophisticated cyber attacks. Several had been observed in other parts of the world that should have served as a harbinger of sorts, but each time they were dismissed with "it could not happen like that in America." Even the insurance industry, wary of such scenarios, did not believe any capable threat actor would attempt a massively damaging attack, let alone succeed.

The countdown to disaster began more than a decade ago, with several cyber campaigns that were discovered and discussed openly in the media, given names like Den of Thieves and Elegant Frost. They showed that tarnished national pride was offended by a series of Western energy and trade policies that left them with a smaller share of the new global prosperity. What infuriated them the most was the prosperity being enjoyed by neighbors, while they began to languish and be outcompeted.

Their own words warned us—they had felt as if they had been pushed into a corner—but what we did not know is that, from the corner, the groups could flip several switches. With the Shanghai, Mexico City, and Bangkok disasters preceding it, the year leading up to what was coined as the worst cyber attack the world had ever seen was uniquely chaotic and dangerous.

There were already several countries that were dealing with a nasty web of insurrection and subversive armed intrusions.

The lessons of hybrid warfare borne out in the 2010–2020 timeframe in eastern Europe were being applied with some effect. It was years of deeply knowing several targeted organizations and their operations that allowed planners to build their plan. The attackers were well positioned, as their country had enjoyed a short season of growth and modernization that brought Western and Chinese firms to upgrade their power systems and cellular networks, and to bring IIoT to their country. These improvements gave the attackers the ability to understand the system of digital bricks that the world was building its future upon. It took over a year to engineer an attack, and months to then position everything in a veil of darkness. Strategic investments in research and technology programs combined a deep technical understanding of modern satellite and atmospheric communication networks, automation and control technology, and a chip-level working knowledge of microcomputer boards.

The attack was prepared and executed in a series of well-synchronized stages.

- *Stage 1.* It all began with a series of implants in meter and microgrid data aggregators and select communication gateways. The code was light weight, easily positioned in a few initial hosts; once in place, it could self-propagate from device to device across the native communication networks. The only trick was to propagate in a manner where the attack did not congest its own pathways and did not get so noisy as to reveal itself in large swaths of traffic where it hit public networks. The simple family of exploits took advantage of an unknown weakness in the code used to enable web-capable management interface. Researchers first published the vulnerability four years prior but no one had put the time into operationalizing a working exploit, or at least no one thought that had happened. The takeover of hundreds of thousands of power grid meters and power inverters provided a large homogenous botnet that could quickly overwhelm New York's telecommunication networks while refusing remote connection attempts by the utility. If you could even get through all the traffic, the devices would no longer recognize authentication attempts. This attack stage created a great deal of confusion while complicating all sorts of

communications that relied on shared networks to receive vital data from the many "micro-processor-based things" that helped the city function.

- *Stage 2.* The second stage was composed of a few select actions to disrupt power flowing in and to one of the world's hungriest load centers. This attack required serious engineering, but once in place, the code would do all the work. Operational traffic captures from a few unmanned substations provided a good look at how the utilities being targeted were applying a common industrial protocol used in SCADA applications. The software implants had been coded to verify the specific implementation of breaker control before it began to send commands to RTUs to open remotely operated circuit breakers and de-energize critical circuits. These precise actions would create pockets of outages pushing the system closer to stability limits. The loss of load would result in an over frequency condition that machines would instantly sense and begin to balance. That is when the final attack would activate.

- *Stage 3.* The last stage of the attack was timed as a final shot before the other malicious codes would turn to a final payload module and overwrite memory at a basic level forcing replacement of the many devices. The long-term prospects of reenergizing the power system that served the city would become bleak in a matter of seconds. The final shot had a 40–50 percent chance of creating a wider outage to be felt outside of NYC. The attackers had been hard at work finding their way onto the operational networks for a number of cloud-connected gas turbines that supplied large portions of the consumed power on the island. Once there they were able to devise two primary methods for placing the turbine in a dangerous condition and after several attempts began to overwrite the system software and firmware. Some of the attacks were successful in changing control setpoints that were able to trip units while a few others actually caused physical damage. The result was a well-synchronized loss of supply pushing the grid back in the other direction. The outage was still contained to the region and manifested itself in several pockets—leaving some microgrid devices and power meters to continue sending a tsunami of messages.

The combination of all three stages overwhelmed grid operators and city managers, creating conditions that stressed the well-practiced plans to deal with all sorts of crises. The city known for its planning and ability to absorb assault was plunged into a dark and an eerie silence. The pause lasted longer than normal as emergency operations personnel waited to see if the power would return and tried to make sense of why they were receiving minimal data from what was recently heralded as one of the most instrumented cities in the world. The optimization cloud applications were providing strange results for fire and police units on their city-wide operational picture displays. Few people knew that the ocean of power meters were jamming networks with constant streams of gibberish packets.

At first everyone focused on the immediate crisis of clogged communications and power outages, but the city's hydrologists knew there were bigger problems to worry about. NYC has been kept dry by a series of huge pumps that remove intruding water into the city's vast underground and returning it to the Hudson. Slight variations in the water height had required a massive city works

project to keep underground vaults dry and allow New Yorkers to use one of the most important services the city offered—public mass transit. The pumps had been configured to receive power from multiple redundant circuits, several key pumping stations were now offline, and the water intrusion spread. Unknown to the attackers, two of the key substations attacked were required to maintain power flow downstream to those pumping stations. The failure of a proper make-and-break configuration on the local backup generator would go unnoticed as alerts were never sent to the city's hydrology ops center. The intruding water set off a number of tiny sensors used to show the spread of water, but that data never found its way to the NYC private cloud providing data to city engineers. The water would actually undermine a valiant effort to restore power to sections of the city as transformers and conduits were energized without knowing sections were underwater. Several electrical shorts occurred, adding to the damage.

It took two days to simply hatch a plan to combat the remaining botnet, and within the first hour the plan would become unnecessary as the meters and inverters began to pop offline, never to reset and reboot and come back. It was not the eye of the data storm as one person joked but the end of the advanced meter network the utility had come to rely on. The utility power meter engineers and security team, analyzing infected meters taken from the field, had missed the module responsible for the firmware overwrite routine as they focused on portion of the code responsible for sending out all the errant messages. The plan would now have to be modified to visit each device and swap them out. New reports were starting to be radioed in or sent via satellite phone that two of the three types of meters had actually performed remote disconnects, interrupting power to homes and buildings. The outage would grow in size for one last time.

The only saving grace was that there were few ways to get to the remaining systems to perform additional cyber attacks. By day three, cell towers, which had only recently been providing sufficient throughput, began blinking off the communications grid while other emergency facilities were suffering from the same fate, losing their back-up generators. The decision to evacuate was a hard one, but no one could provide a confident estimate for restoring power at the edge of the system, where meters had been bricked. Even worse, the city was literally flooding from the basements up, making habitation a health risk and further undermining efforts to move and care for people. The mayor requested the governor send in the National Guard to help utility personnel remove meters for direct connections. The procedure was not complex but it did require two-man teams to visit hundreds of thousands of locations. A return to normal would be measured in neither days nor weeks, but more likely months if not years. And it would certainly have to be a "new normal."

Conclusions and Recommendations

In 2017, many of the current IoT products are literally toys—some, like Wi-fi-enabled Hello Barbie, are intended for children. Others, like increasingly capable drones, and IIoT products, like highly connected industrial equipment, are obviously made for adults. Ubiquitous cloud-computing and storage capabilities are already in wide use by children and adults to such a pervasive extent that

much of our modern world—from households to businesses to industrially intensive operations—would cease to function if disconnected from cloud services for any appreciable length of time.

Now mix in the accelerating rate at which connectivity between not just intelligent objects and the cloud, but between objects and other objects, is expanding, and the degree of interdependence we're building and accepting is simply staggering. Boy do things work great when they work. But what are our plans B and C for when we these things fail or the systems on which they depend fail? And fail they will.

Below find a starter list of cautionary observations, each of which suggests its high-level solution.

- Cell overload—The technologies implemented to digitize infrastructures have outpaced the cell networks they relied upon.
- Restoration overload—Large deployments of things (e.g., instruments/sensors) can quickly outpace stakeholders' ability to maintain or restore them if a widespread common failure or attack takes place
- Mass cascades—Single system disruptions can quickly cascade as large number of people's routines or plans are changed resulting in capacity surges and difficult-to-predict impacts as first-order impacts are accompanied by second- and third-order impacts
- Software at scale—Software introduces large-scale systemic risk when implemented in large scales. Updates need not only to be tested and receive device-level quality checks, but system-wide modeling or simulation is necessary
- Software defects—Software-induced errors can serve as a blueprint for a malicious attack if access to maintenance or engineering systems can be obtained
- Mono-culture risk
- Difficulties in risk detection because of complexity, opacity, latency, and disguise

What's it going to take to follow through on any of these suggestions? Accidents, property damage, corporate reputational damage, national security impacts, injuries, and significant loss of human life. In short, problems that individuals, companies, and governments recognize today as safety problems. Security pundits, particularly those focused on cyber security risks to industrial operations, have been warning for years that interconnecting and automating systems that control often-highly dangerous physical processes brings with it a type and scale of risk we had previously not seen. Many have said that the answer lies in fusing security matters with safety culture.

We've seen cars, phone, toys, and many other types of tech-enabled products recalled or terminated due to safety issues. When the same business and social impulses begin to extend into the security realm, when more industrial software has to meet the requirements of "safety critical" systems, we may find ways to avoid the scenes such as those depicted in this paper.

1616 RHODE ISLAND AVENUE NW
WASHINGTON, DC 20036
TEL. (202) 887.0200
FAX (202) 775.3199

Our civilization is grappling with unbounded complexity and cyber exposure brought by automating important processes without a full consideration of the possible cyber consequences. Obvious and seemingly unstoppable trend lines are pointing to massive deployment of increasingly automated and even autonomous systems underway now and accelerating over the next few years. We recommend a strategic pause to reconsider how we more fully value automation from a cyber-informed cost-benefit perspective. And with or without that pause (we assume most won't understand the rationale) it is imperative that we find ways to identify, interrupt, and prevent catastrophic cyber-physical consequences of both cyber-attack and malfunction of these technologies.

Acknowledgments

This report is made possible by general support to CSIS. No direct sponsorship contributed to its publication.

About the Authors

Michael Assante is a senior associate with the Technology Policy Program at the Center for Strategic and International Studies in Washington, D.C. Andrew Bochman is senior cyber and energy security strategist for the Idaho National Laboratory (INL) in Idaho Falls, Idaho.
