Spring 2013 – Machine Reputation – Is Data Trustworthy?

Introduction

In this paper, I will show some of my impressions of machine reputation. The idea that data from machines has more value than the same information from humans seems interesting and disturbing. This paper will continue with some key point discussions from sources to help convey these ideas, and to provide the reader resources for further research. Lastly, I will conclude the paper with a brief descript of some final thoughts on the subject. .

Bias from impartial bits – Assumptions of machines in medicine and law

Several years ago I was watching TV and stumbled upon a channel that was already running a program. The program was a documentary and it showed a doctor followed by several medical students at the bed of an incapacitated patient. The narrator was pointing out that the patient was attempting to get the doctors attention because they were too warm. The doctor paused with his students and checked the machines attached to the patient. He then reassured the patient that their temperature was normal and continued on with his students.

I thought it interesting at the time since my work also required me to take personal accounts into consideration when dealing with technical issues. I had experienced this same thing, with my role as the doctor, and it was alarming to think that a bias of reputation was already established. There is no place where the implication of bias is greater than a court of law.

"When human witnesses take the stand, the triers-of-fact are expected to generally consider the possibility that they, despite the oath, may render an untruthful or factually incorrect account of events and circumstances due to a conflict of interest or bias."[1]

"However, when computer-generated data is introduced as evidence in court, there appears to be a strong assumption that such evidence is somehow impartial and as such more trustworthy than testimony given by a human witness or an expert witness."[1]

Since the assumption that machine data is accurate for reasons of validity, the scrutiny of the same data is countered based on this bias. "One important consideration in such cases that involve transient events captured only by a single instance of software (and all the more so when the software is for the plaintiff) is that the defendants are foreclosed from exonerating themselves by providing independent sources of evidence or causing independent tests to be performed (such as with exonerating DNA evidence)."[1]

There are several problems with the steadfast reliance on machine accuracy, in which a machine is dependent on its programing code. First, machine code is subject to faults. The code may be entered incorrectly by humans for reasons of malice, fatigue, or lack of skill. Next, the program code may not be designed for the intended use that it was applied to, resulting in data that is outside the range of acceptable tolerances. Lastly, program code can be influenced by outside factors, such as sabotage, incompatibility, or system dependencies that are outside the control of the program code. Unfortunately, this data can be used in civil court cases that don't provide the same protections under law[2].

The road to hell is paved with good intent – When things don't go as planned

Computation systems, sensor networks, and automation are heading into an unknown frontier. There are some speculations as to what that frontier will be like, but no one is absolutely sure. The serious implications are a matter of priority. This has opened debate between nations, as well as institutions, government, and scholars.

It is safe to assume that we are now dependent on the machines we have created to accomplish our goals and solve our problems. Fine tolerances and resolution have been brought about by the subsequent use of machines. Without this level of precision, much of the work done today simply could not be done.

One example of everyday use of digital machines is the automobile. The OBD (On Board Diagnostic) Program was brought about by California legislation in response to poor air quality in the state. The state's EPA was responsible for the effort of research to the causes and the solutions to correcting the problem. The state mandated that any car manufactured after 1996, and sold in the state, shall have the OBD II (second generation) system. This not only allowed EPA station inspectors the ability to quickly determine if a vehicle is within emissions compliance, but it also allowed the vehicle to monitor and control its emissions automatically. [3]

The OBD system contains at least one ECU (Engine Control Unit), OBD connector ports, a host of sensors throughout the vehicle, and actuators that respond to the ECU commands. The premise of the OBD system was to allow the car to take corrective measures during events that would create more pollution. If it was unable to make the adjustments, it would notify the driver of a fault, usually with a "check engine" light. The system was designed with the assumption that no human would be required to interact. The only exception to this would be a state inspector or service technician that would reference the operation of the vehicle and to reset any alarms.

The University of Washington and University of California, San Diego conducted a research project into security of the OBD system which was a first of its kind. The study was released to the public at the IEEE Symposium on Security and Privacy 14 years after the California mandate. The findings were revealing.

The researchers pointed out that "Modern automobiles are no longer mere mechanical devices; they are pervasively monitored and controlled by dozens of digital computers coordinated via internal vehicular networks." They also pointed out that "Outside the academic realm, there is a small but vibrant "tuner" subculture of automobile enthusiasts who employ specialized software to improve performance (e.g., by removing electronic RPM limitations or changing spark timings, fuel ignition parameters, or valve timings) frequently at the expense of regulatory compliance"[4]

This innocent rogue element was not considered during the development and implementation of the ODB system. Without this factor being in the fold, the behavior of the entire OBD system can be called into question. Up to this point, the rogue enthusiast is an amusing novelty. However the researchers demonstrate that their findings were anything but amusing.

"We demonstrate that an attacker who is able to infiltrate virtually any Electronic Control Unit (ECU) can leverage this ability to completely circumvent a broad array of safety-critical systems."[4]

"The first is physical access. Someone-such as a mechanic, a valet, a person who rents a car, an ex-friend, a disgruntled family member, or the car owner-can, with even momentary access to the vehicle, (can) insert a malicious component into a car's internal network via the ubiquitous OBD-II port (typically under the dash)...The other vector is via the numerous wireless interfaces implemented in the modern automobile." [4]

"Control of the BCM's function is split across the low-speed and high-speed buses. By reverse engineering...we were able to control essentially all of the BCM's functions...lock and unlock the doors; jam the door locks...pop the trunk; adjust interior and exterior lighting levels; honk the horn (indefinitely and at varying frequencies); disable and enable the window relays; disable and enable the windshield wipers; continuously shoot windshield fluid; and disable the key lock relay to lock the key in the ignition...Additionally, we can forge a packet with the "airbag deployed" bit set to disable the engine." [4]

"A more sophisticated attack could implant malicious code within the telematics environment itself (either in RAM or by re-flashing the unit). Doing so would allow the malicious code to co-exist with the existing telematics software"[4]

"Finally, we also found that, in addition to being able to load custom code onto an ECU via the CAN network, it is straightforward to design this code to completely erase any evidence of itself after executing its attack. Thus, absent

any such forensic trail, it may be infeasible to determine if a particular crash is caused by an attack or not. While a seemingly minor point, we believe that this is in fact a very dangerous capability as it minimizes the possibility of any law enforcement action that might deter individuals from using such attacks."[4]

The reports mention of code that co-exists with the existing software with the ability to hide itself was eerily familiar with another set of circumstances brought to light around that same time, Stuxnet.

He who has the gold makes the rules – When one plus one no longer equals two

"A sophisticated worm designed to steal industrial secrets and disrupt operations has infected at least 14 plants, according to Siemens. Called Stuxnet, the worm was discovered in July (2010) when researchers at VirusBlokAda found it on computers in Iran."[5]

The Stuxnet worm was unique because it didn't do what most viruses typically do, which is attack a computer that a human uses. The target wasn't random, it was the Siemens PLC SCADA industrial control system. These systems are used for various industrial applications. The particular use of the infected systems was even more unique. These systems were used by the Iranian government to control centrifuges for enriching uranium. Those that identified the worm speculated that the origin of the worm was a state sponsored effort. Symantec's O'Murchu agrees that the worm was done by particularly sophisticated attackers. "This is definitely not your typical operation," [5]

It was the length to which the worm was designed that led those that identified it as too "sophisticated" for a private enterprise to develop. The methods that were used to propagate the worm were high level.

"All, the functionality required to sabotage a system was embedded directly in the Stuxnet executable. Updates to this executable would be propagated throughout the facility through a peer-to-peer method established by Stuxnet."[6]

"One of the main propagation methods Stuxnet uses is to copy itself to inserted removable drives. Industrial control systems are commonly programmed by a Windows computer that is non-networked and operators often exchange data with other computers using removable drives. Stuxnet used two methods to spread to and from removable drives—one method using a vulnerability that allowed auto-execution when viewing the removable drive and the other using an autorun.inf file." [6]

One of the most interesting aspects of how the worm worked was by means of a corrupt device driver signature. The reason for this is that digital signatures are an agreed upon trust, established by an international collaboration of government, industry, and academia. The attackers had managed to compromise the digital signature used by Realtek JMicron which is assigned by Verisign, the authority that the great collaboration had tasked with its trust.

The manufacturing of atomic weapons by the Iranian government is clearly an international concern. Any proliferation of weapons should be discouraged. The debate of this is outside the scope of this paper. However, the trust in digital systems that are agreed upon as being true is the focus of this paper. That trust was clearly compromised and cannot be deemed reliable.

The sum of all parts – Machines mirror the humans that make them

> *"It can only be attributable to human error." - 2001: A Space Odyssey (1968 - Stanley Kubrick and Arthur C. Clarke)*
>
> *"We aren't dealing with ordinary machines here. These are highly complicated pieces of equipment, almost as complicated as living organisms. In some cases, they have been designed by other computers. We don't know exactly how they work." - Westworld (1973 - Michael Crichton)*

Even with all precautions in place, there will remain the possibility of something unexpected. Even seasoned developers still experience surprises. These "glitches" can range from an annoyance to a tragedy.

"First Blood" occurred when Matt's quadcopter flew into his face on Christmas Eve. Though the firmware accepted an arming sequence, the radio must have had some configuration issue on the throttle channel (exponential throws, narrow throws, etc) that caused this problem."..."Matt took 11 stitches to the face and was back at work after Christmas Day."[7]

One account of a failure was brought about by a phenomena know in the metallurgy community as metal whiskering or tin whiskers. This is a problem with far reaching consequences. The subject was discussed during the 5th International Tin Whisker Symposium by NASA staff.[8]

In it they make mention of issues of unintended acceleration resulting from this phenomena with direct references to faults with the 2005 Toyota Camry. The presentation goes on to show which senor had failed, causing the undesirable acceleration, as well as a detail view of the actual short circuit. They also demonstrate the difficulty of visual detection by comparing the width of human hair in contrast with the whisker, one measurement was shown to equate to 1.1um. The width of a healthy red blood cell is 6-8 times that of the tin whisker measured.

To make matters worse, the Toyota Motor Company's quality control had no means to prevent this, even if they had sophisticated sensing equipment. The reason for this is the whiskers grow from the metallic surface over time. The issue is further exacerbated by the strict rules of "lead-free" electronics. When lead is part of the solder alloy, it actually inhibits the formation of whisker material. This particularly simple problem cost the lives of several people.

The details of some malfunctions have conflicting reports and lack any formal report available to the public. This is particularly true with defense systems, which require extensive secrecy. The operation of defense systems is largely guarded. Any errors that do occur have the potential for revealing vulnerabilities that have a direct consequence of diminishing national security.

"In April 2008, several TALON SWORDS units—mobile robots armed with machine guns—in Iraq were reported to be grounded for reasons not fully disclosed, though early reports claim the robots, without being commanded to, trained their guns on 'friendly' soldiers [9]; and later reports denied this account but admitted there had been malfunctions during the development and testing phase prior to deployment [10]. The full story does not appear to have yet emerged, but either way, the incident underscores the public's anxiety"[11]

The faith of public safety is based on proper design, adherence to regulations, and thorough testing over an extended period of time. Unfortunately these systems, and the resources to develop and deploy them, have limits. It often becomes more of a balance of risk and mitigation, with the development sandwiched between costs and returns.

"October 2007, a semi-autonomous robotic cannon deployed by the South African army malfunctioned, killing nine 'friendly' soldiers and wounding 14 others [12]." The risk of damage, injury, or death are so high for defense systems, the margin of error is simply unforgiving. One hypothesized, but serious example of the potential implications of automated defense system failure would be the death of US citizens. This example does not suggest that such an event had occurred, but only to emphasize the gravity of an unforeseen error.

In 1996, TWA flight 800 from New York's JFK International airport exploded and crashed 12 minutes after takeoff, killing all 230 people onboard. Several eyewitness accounts suggest that the accident was a result of something originating from the surface and striking the plane. The investigation was extended beyond a technical scope, handled by the NTSB, and into a criminal scope, handled by the FBI. The investigation was concluded four years later. The report stated that the accident resulted from an electrical short circuit in the fuel tank, which caused an explosion and compromised the structural integrity of the air frame.

On June 19th, 2013, a new documentary claims the final report has flaws and that key evidence had been dismissed, based on testimony from investigators of the incident. Whether these claims are true or not, the US Navy did have an automated system for engaging enemy at the time of the accident.

"The US Navy's MK 15 Phalanx Close-In Weapons System, the earliest model, was first installed on a ship in 1980, and modified versions are still widely used by the United States and its allies. It is designed to sense approaching anti-ship missiles or threatening aircraft and respond with fire from two 20mm guns with six rotating barrels. The

guns each fire 3,000 to 4,500 rounds per minute. More recent models aim to defend against small gunboats, artillery, and helicopters. The Navy describes the Phalanx as "the only deployed close-in weapon system capable of autonomously performing its own search, detect, evaluation, track, engage and kill assessment functions."[13]

The thought that this accident was factually a result of an autonomous system is truly dreadful. With that said, the point that needs to made clear is any system with the potential for such dire consequences has the greatest burden of reliability. In the absence of this event, a proven application of failure mitigation must be executed to prevent the potential of tragedy because the consequence is far too great.

Summary

"So far the predictions about what (machines) will be able to do has been consistently over-optimistic. Whatever the length of time required for those predictions to come true it is already obvious that man has little to fear, but that he could well fear a future without them." [14]

Is it safe to assume that the experts are correct and we have nothing to fear? The answer appears to depend on what the complexity, design, and openness of the technology are. It also seems to depend on who you ask.

A doctor or a lawyer most likely will leverage their professional experience over any symptom or evidence brought before them. These professions have become reliant on computation systems and the data they provide. This has created a professional bias of reliability without any qualified understanding of the computer systems.

The pervasive nature of computation systems has created a comfortable culture that is unaware of its presence. It is difficult to find an entire day that a human isn't interacting with some kind of microcontroller that senses and responds. Systems with intended uses can ultimately be placed in situations that they were not intended be. The complexity of the systems can also overshadow the risks facing the public.

Some technologies contain several layers of interconnect subsets that could be compromised in any part results in a global failure. Without segmented understanding, these faults may not be recognized until some detectable condition merits further investigation. Failures on critical system can have far reaching consequences. Given the nature of a fault, some outcomes could fall outside the realm of ethical comprehension.

Conclusion

The shortcomings and failures pointed out in this paper only reflect an extremely narrow range of possible problems. In all, these faults are an extension of the burdens that confront all people. Mistakes will be made and there will be consequences for those mistakes. No certainty exists in what we create, or the changes that will result after it is created.

Is data trustworthy? It depends on the full understanding of the source. It also relies on the validation and predictable response from an analogous source to ensure that the result is indeed true. Without any assurances, the data is suspect.

Sources

[1] - Software on the Witness Stand: What Should It Take for Us to Trust It? - (Dartmouth College) Authors - Sergey Bratus1, Ashlyn Lembree2, Anna Shubina1

[2] - "In all criminal prosecutions, the accused shall enjoy the right . . . to be confronted with the witnesses against him." U.S. Const., Amend. VI.

[3] - California Environmental Protection Agency - On-Board Diagnostics (OBD) Program
http://www.arb.ca.gov/msprog/obdprog/obdprog.htm#background

[4] - Experimental Security Analysis of a Modern Automobile - 2010 IEEE Symposium on Security and Privacy - Karl Koscher, Alexei Czeskis, Franziska Roesner, Shwetak Patel, and Tadayoshi Kohno - Department of Computer Science and Engineering - University of Washington
Stephen Checkoway, Damon McCoy, Brian Kantor, Danny Anderson, Hovav Shacham, and Stefan Savage - Department of Computer Science and Engineering - University of California San Diego

[5] - Computerworld - Siemens: Stuxnet worm hit industrial systems - Robert McMillan (September 14, 2010)

[6] - Symantec W32.Stuxnet Dossier, Attach Scenario - (Feb 2011) Authors Nicolas Falliere, Liam O Murchu, and Eric Chien

[7] - A Year of Quadcopter Experimentation (2013 - Ken Gracey, President, Parallax)

[8] - "Electrical Failure of an Accelerator Pedal Position Sensor Caused by a Tin Whisker and Discussion of InvestigativeTechniques Used for Whisker Detection". 5th International Tin Whisker Symposium by NASA staff Henning Leidecker, Lyudmyla Panashchenko, and Jay Brusse. September 14th, 2011

[9] - Page, Lewis (2008). "US War Robots 'Turned Guns' on Fleshy Comrades", The Register (UK), April 11, 2008. Last accessed on September 15, 2008: http://www.theregister.co.uk/2008/04/11/us_war_robot_rebellion_iraq/

[10] - Sofge, Erik (2008). "The Inside Story of the SWORDS Armed Robot 'Pullout' in Iraq: Update", PopularMechanics.com, April 15, 2008. Last accessed on September 15, 2008: http://www.popularmechanics.com/blogs/technology_news/4258963.html

[11] - "Autonomous Military Robotics: Risk, Ethics, and Design" (December 20, 2008) Patrick Lin, Ph.D., George Bekey, Ph.D., Keith Abney, M.A. - California Polytechnic State University, San Luis Obispo

[12] - Shachtman, Noah (2007). "Robot Cannon Kills 9, Wounds 14", Wired.com, October 18, 2007. Last accessed on September 15, 2008: http://blog.wired.com/defense/2007/10/robot-cannon-ki.html

[13] - Losing Humanity - The Case against Killer Robots - Human Rights Watch (NOVEMBER 2012 ISBN: 1-56432-964-X)

[14] - Robots (1978 – Jasia Reichardt)