Innovation of Cyber Defense Technologies with Data Analysis and Collection

CS4991 Capstone Report, 2023

Michael Kosar
Computer Science
The University of Virginia
School of Engineering and Applied Science
Charlottesville, Virginia USA
mjk5pt@virginia.edu

ABSTRACT

The United States Air Force decided to design a tool to differentiate between different radio interference signals in the X-band due to the increasing threat of cyber warfare and adversarial communications disruptions. In order to complete this project, I used various data collection and analysis tools to develop a Python computer program that was able to parse and classify large sets of data into usable charts. Using the Python programming language and various libraries such as NumPy, I wrote a program to comb through large sets of real-world data and classify it based on various factors such as length of transmission, altitude, and geographic location. By filtering out the unwanted data, and sorting the necessary data, the program was then able to use multiple algorithms to put together a physical chart index categorizing the threat level of the interference based on the real-time information gathered. The end product was a chart with five different threat level tiers that can be deployed and used by the government in real world scenarios. It provides U.S. warfighters with a fast and simple way to identify adversarial cyber threats based on the data collected so that they can respond with appropriate defensive measures. In the future, more time would be needed to do extensive testing in order to maximize accuracy in determining adversarial threats and ensure the chart is up to date as threats become more advanced.

1. INTRODUCTION

The adoption of technology as the leading source of data storage, the modernization of nations, and the increasing reliance on technology to complete daily tasks has led to an increase in cyber vulnerabilities and cybercrime worldwide. There has been much talk about the rise in cyberattacks within the last decade and how these attacks have affected the victim in each situation. Victims can include governments, large corporations, small businesses, or even families. The only time these attacks become mainstream news is when they affect the day to day lives of the public. Such an example would be the recent Colonial Pipeline cyberattack in which a group of hackers took control of the largest refined oil pipeline in the United States and held it for ransom. This left much of the east coast of the U.S. without gasoline for days and subsequently caused an uproar about the security of our cyber infrastructure.

But this is just one side of the cyber warfare going on around the globe. There is another, much more serious, and much less talked about side that is being fought by the U.S. military. The United States is constantly defending itself against cyberattacks from foreign adversaries around the globe. So much of the U.S. infrastructure, data storage, and technologies are connected to digital systems. This makes them a prime target for foreign groups and countries that wish to do harm to

the U.S. or steal its secrets. Because of this, a mini arms race has risen between the U.S. and its enemies as each wishes to build, design, and implement better and more sophisticated technologies in order to defend against and attack each other. There is constant pressure to thwart the next attack and develop capabilities that, if the day came, could infiltrate or shut down an enemies digital network. Among the groups in the United States military involved, one of the main targets is the U.S. Air Force.

In the event of a war, the United States Air Force would be one of the first entities to engage in the fighting. They have outfitted their aircraft with the most up-to-date and sophisticated radars, sensors, and technologies that are available. The majority of the U.S. Air Force's technologies rely on radar and transmission data. Typical aircraft have around 180 touch points across several networks as well as a variety of onboard processors (Maybury, 2015). This makes them prone to cyberattacks and interference from adversaries. For this reason, they are putting a large effort into securing their systems and finding ways to identify threats before they occur. One of the main ways that adversaries can disrupt communications and aircraft systems is with jamming. This is when radio-waves are sent to disrupt certain technologies that also use different radio waves and frequencies to operate. In order to stop this, the Air Force needed a way to distinguish between harmless, everyday communications that are constantly flowing through the air, and malicious, adversarial waves that could potentially cause harm to the aircraft or pilot.

2. RELATED WORK

In order to gain a deeper understanding in the field of radar equipment and signal processing, I frequently referred to Richards (2014), who covered basic radar topics such as the different types of radars, what they are used for, and how each type operates. This allowed me to get a deeper understanding about how radar technologies work so that in developing my tool, I could understand how these radar systems were operating. It allowed me to also understand the capabilities of certain radars and what they can and cannot do. This helped me develop an approach that was feasible and not too difficult to implement.

Multiple times throughout this project I referenced and used the NumPy library documentation, which highlights the uses and strategies of the NumPy python library. This built-in library in python code allows the user to easily manipulate and sort data through prebuilt functions and tools. The documentation allowed me to better understand the tools at my disposal and how to use them properly. This not only made my work a lot easier but allowed me to complete it more efficiently and output a better result through prebuilt tools.

3. PROCESS DESIGN

The first step in this project was to look at the raw data given to us and determine the best way to analyze, organize, and implement it. We were given large quantities of data sets containing radio wave information in the X-band that was gathered from fighter jets in real test scenarios. The data contained information pertaining to the frequency, wavelength, altitude, length of transmission, and geographic location of the waves. In order to break the data down into usable sets, we created a Python program that used the NumPy library to sort the data based on the information we determined to be most relevant. This program organized the data in order of magnitude for frequency, wavelength, and length of transmission. We decided to leave out geographic location as we did not believe this to be a necessary data point. We also did not include the altitude data as transmissions can be found at all altitudes that fighter jets perform at due to the constant

communication between ground and satellite links.

Once the data was organized, we needed to define our parameters and values to input into our sorting algorithm. We first looked at the frequency data points to understand the commonalities and ranges. We were strictly working in the X-band because this is the band that the U.S. military typically operates in. The typical frequency range for the X-band is considered to be 8.0 - 12.0 GHz. The wavelength in the band is generally from 3.75 - 2.5 cm (NASA). Because we wanted to stay in this band of frequency, any data gathered outside of this range was discarded. The Radio Regulations of the International Telecommunication Union allow amateur radio and satellite communications in a frequency range from 10.0 to 10.5 GHz. Because this range is typically used for amateur communications, we decided to classify the data in it as the lowest risk of interference. Enemies tend to jam radars using high frequencies so we decided that higher frequencies, from 10.6-12 GHz, would be considered more dangerous. The lower frequency range, from 8.0 - 9.9 GHz, would be considered potentially dangerous because lower frequencies have higher wavelengths.

The next factor considered is the wavelength. Higher wavelengths have more penetrating power to go through obstacles as well as travel farther than smaller wavelengths. We can then assume that they are the ideal length used by ground-based disruption systems to target something far away. Because of this, the longer wavelengths, 3-4 cm, were classified as more dangerous than the shorter ones. Shorter wavelengths are typically used for high resolution imagery (Steur, 2018) so although they may be used for spying or surveillance, they are not a high active threat of disruption.

The final factor we took into consideration was the length of transmission. The longer the transmission lasted, the more likely it was intended to disrupt communication. This is because as an aircraft is traveling through the air, it may briefly encounter and interfere with normal transmissions. Shorter bursts of transmissions are not as harmful or disrupting so they were classified lower on our scale. But if a transmission was constantly interfering with the system, this meant the aircraft was most likely being targeted.

We used a sorting algorithm to group together the various data points based on our classifications and their order of importance. We decided the most important factor was frequency, followed by wavelength, followed by length of transmission. The data was then sorted on the basis of those three factors, in order of most to least dangerous. For example, data points with high frequency, long transmission lengths, and high wavelengths were ranked higher than data low frequency, short transmission lengths, and low wavelengths. Dividing this sorted data into 5 equally spaced categories allowed us to develop our final chart.

4. RESULTS

The data at the top of the list was considered to be in tier 5. This was the highest tier, indicating an active threat that must be either defended against or avoided. This tier generally had a frequency of 11-12 GHz, a wavelength of 3.5-3.75 cm, and a transmission length longer than 1 minute. Level 4 indicated a likely threat that could possibly disrupt communications. This tier generally had a frequency of 10-11 GHz, a wavelength of 3.25-3.5 cm, and a transmission length between 45-60 seconds. Level 3 indicated that the wave could be either passive or adversarial and for which the user should exercise increased caution. This tier generally had a frequency of 9-10 GHz, a wavelength of 3.0-3.25 cm, and a transmission length between 30-45 seconds. Level 2 was the second lowest threat level, indicating a mostly passive radio

wave that had a few characteristics of a potentially adversarial wave. This tier generally had a frequency of 8.5-9 GHz, a wavelength of 2.75-3.0 cm, and a transmission length between 15-30 seconds. Level 1 was the lowest threat level, usually indicating a passive, common transmission such as a civilian radio communication. This tier generally had a frequency of 8-8.5 GHz, a wavelength of 2.5-2.75 cm, and a transmission length less than 15 seconds.

Although there is not much known regarding how this chart is being implemented currently, it allows aircraft to quickly identify incoming transmissions as either passive or aggressive. When the aircraft intercepts a transmission, it can automatically input the data gathered from it into a program that uses our classification chart. The program compares the data to the chart and outputs a message defining the potential severity of the transmission. This could help change the battlefield and give fighters the ability to actively defend against aggressive attacks.

5. CONCLUSION

The growing threat of cyberattacks on the United States military has led to the large allocation of resources towards its defense. Specifically, there has been an increased focus within the U.S. Air Force to enhance its cyber defense capabilities. Using data collected from real-world test scenarios, I helped develop a tool that could be used to defend against radio-wave interference in the X-band. Through the use of data collection and analysis techniques, along with my knowledge of computer programming and algorithms, I was able to sift through large data sets and turn them into an organized chart. This chart could potentially give the U.S. Air Force an edge on the battlefield through the quick detection, analysis, and categorization of incoming radio-waves. By helping to determine whether the waves are harmless or adversarial, this tool will hopefully be able to make a meaningful difference in the future to protect its users.

6. FUTURE WORK

The next steps in this project would be to test this technology in training scenarios. Due to time limitations, I was not able to properly test the effectiveness of this chart. By allowing the use of this technology in real-time scenarios we can properly test the accuracy and effectiveness of it. The additional data gathered would allow us to better refine the data sets and ensure the correct results. A broader range of data gathered would be beneficial to the chart as well. An example of an area that would be important to analyze is reaction and processing time. In a real scenario, the processing time between detection of a signal and analysis of the threat level would be critical to the effectiveness of the response. In the future, I would gather larger data sets, broader data points such as reaction and processing speed, and test the technology in order to refine its accuracy.

REFERENCES

Maybury, M. (2015). Toward the assured cyberspace advantage: Air force cyber vision 2025. *IEEE Security & Privacy, 13*(1), 49-56. doi:10.1109/msp.2013.135

National Aeronautics and Space Association. (n.d.). *Basics of space flight - solar system exploration: NASA science*. NASA. Retrieved March 15, 2023, from https://solarsystem.nasa.gov/basics/chapter6-3/

Richards, M. A. (2014). *Fundamentals of Radar Signal Processing*. McGraw Hill.

Steur, E. (2018, June 20). Shorter wavelengths for better imaging. Nature Bioengineering Community. Retrieved March 15, 2023, from https://bioengineeringcommunity.nature.com/

posts/29148-shorter-wavelengths-for-better-imaging