Prospectus

An extension in the Chrome Browser for the UVA Library

(Technical Topic)

Artificial Intelligence in Cyber Security

(STS Topic)

By

Nitesh Parajuli

10/30/2019

Technical Project Team Members: Ryan Kelly, Ben Ormond, Tho Ngyuen, Nitesh Parajuli, Yukesh Sitoula, Ben Spector, Ashish Upadhayaya

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: Nitesh

Approved: _____ Date _____
Rider Foley, Department of Engineering and Society

Approved: Ahmed Ibrahim Date 11/26/2019
Ahmed Ibrahim, Department of Computer Science

Technical Introduction

The UVA Library has an abundant amount of resources for different needs, databases for research, the Virgo website to access the library's books, articles, etc. and forefront technologies like 3d printing and virtual reality headsets. While these resources are easily accessible to students, they are also easily ignored. According to the survey conducted by UVA, both graduate and undergraduate students indicated that around 35-40% occasionally use physical library materials, and 33-37% never even use it (cite both). The platform to provide these resources is also showing a change, as shown by the survey that 55% of graduates use online resources weekly, and about 31% of undergraduates used them occasionally. The usage of physical library resources is declining, and the use of online materials is increasing.

Students are inclined to using online resources to find textbook content because the cost of textbooks is too high. According to the National Association of College Stores, the average student is likely to spend around $655 every year (Kingkade, 2017). A textbook trend research showed that about 63% of students look online for book's content. The trend also showed that about 57% of students look online or at other retailers to purchase textbooks (Harris, 2018). The cost of providing library materials itself is demanding, like the UVA's library, which has an expenditure of $36.1 million on providing students with valuable resources (Quick Facts, 2017). These resources will go to waste because there will be a lack of students interested in accessing these resources. Students can access the library materials for free of use, which is a better alternative than spending money on expensive textbooks. In this manner, students can find free resources, and also the library's resources do not get wasted. However, there is still the problem of how students will access the library's resources with ease. A solution is to streamline the resources to students through online software.

To address this issue, a group of six other people and I are building a chrome extension for the Google browser. Since google browser has most of the market share and is the most used browser, it is currently the base for this chrome extension (NetAppliations, 2016). Now, the extension runs whenever a user finishes searching for a book on Google scholar, Barnes & Noble, and Amazon. After the search, the extension sends a popup with the information about the book, such as title and author. While this is a work in progress, in the future, it will also include the availability of the book, location of the library, and login functionalities for students.

Virgo, the UVA Library database, hosts many different forms of resources like books, articles, archives, videos, and many more resources. The Virgo website is currently accessible to everyone. It provides access to reliable and trustworthy materials, while on the other hand, the Internet could give false information. Since Amazon is the largest e-commerce provider, it is also the most used shopping tool, and books are also part of their products. Barnes and Noble is also the largest bookseller (Farfan, 2019). The extension will run on both of these company's website to provide students with alternative options like books for free of costs. Also, when people rely on online tools like Amazon or Barnes and nobles for books, it eliminates the need for a robust technology like Virgo. The chrome extension can be the middle ground for students and researchers to use Virgo.

Technical Topic

The central aspect of the web extension tool is to provide ease of access to the students and other users. One benefit to notice from this extension is that it will easily integrate with the user's search functions. There are no extra steps to be taken, such as traversing through the library website and then searching through their database and then finding the result. The other

main reason this tool is invaluable because it is free of cost. Most people look for books through common websites, and this extension is built to work on those websites. The extension finds similar books in the UVA library and then tells the user about the location and its availability through the popup, as shown below in Figure 1. Again, this is a work in progress. Therefore, some instances are missing.

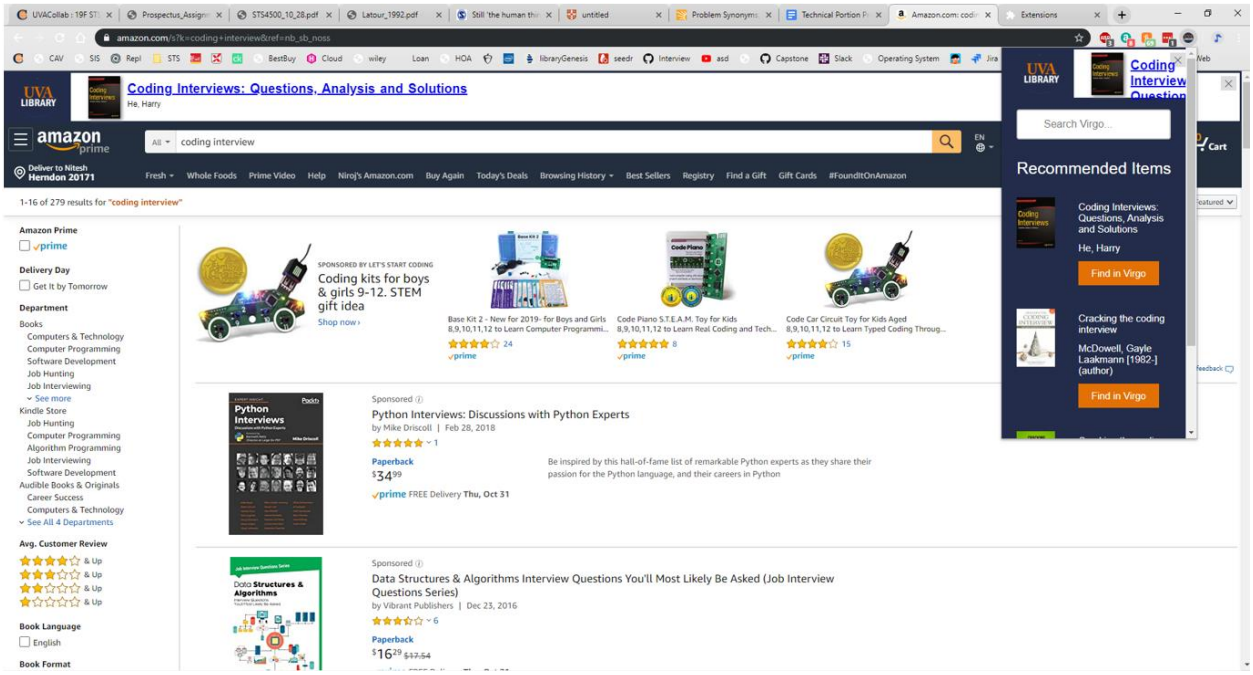

Figure 1. The picture shows the resulting page after searching for a book in amazon. (Parajuli, 2019)

Currently, the extension itself only works through the UVA Library, but in the future, this extension can have the potential to scale. Other universities, high schools, and public libraries can also use this tool. The limitations of this extension are not bound to books as well. It can scale into a research tool as well, helping students and researchers finding journals in their subject of research with ease.

The extension itself is almost close to a minimally viable product, and some functionalities still need to be addressed. The extension will also show students that 3d printers or virtual gears also available at appropriate locations for free to use. The UVA library has also

communicated with local libraries to do cross search for resources. This will promote the Inter Library Loan system the UVA has in place already. The plan is to finish this product by the end of this semester. Testing the product will be the main focus for the next semester.

Most students are not entirely using the resources available to them through then UVA library. They rely on popular search engines or resource providers but at the cost of paying for those resources. To alleviate this problem and help students find an extensive pool of reliable resources, a group of six people and I are working on a chrome extension. The extension's primary goal is providing ease of access to UVA students in finding resources for free. The extension easily integrates within their normal search functions. The extension also has to potential to scale for other institutions to implement similar functionalities.

AI in Cyber Security

The technical aspect focuses on building a chrome extension for the UVA Library while the research will focus on implementation of AI applications in Cyber security. Cyber-attacks has been increasing at an alarming rate. Human intervention is not nearly enough against mutable attacks like computer worms and viruses. According to Dilek, Cakir, and Ayudin (2015), humans cannot respond in time to form an analysis of the attack and develop an adequate response plan. The amount of information a cyber officer needs to analyze is massive, and analyzing it is one thing, but responding to an attack in time is another problem. These tasks are also tedious and repetitive. A solution to solving repetitive labor is already present. The use of computer technologies that automates manual labor. Therefore, cybersecurity also needs a stronger solution. The most promising solution is Artificial Intelligence, but it is not fully integrated within cybersecurity.

To understand why AI is even needed, a network theory is built to realize each actant and how actions of each actant lead to the conclusion that AI can solidify the cyber defense. The actor-network theory was introduced by Bruno Latour, to understand the different combination of elements or actants and interactions between them (Latour, 1992). The actants include humans like cybersecurity officers, Internet users, and hackers. It also includes non-human actants like organizations fighting cybercrimes, organizations that are instigating cybercrimes, software to prevent and detect crimes, and software used with malicious intents. The actants also include ideas like the different applications of Artificial intelligence to combat cybercrimes, cyberthreats, and the impact of intelligent technology influencing people's decision. Enforcing Artificial Intelligence to deal with cybercrimes is based on the fact that cybersecurity officer cannot do enough to battle these crimes and the shape of the cybercrimes are drastically changing at a pace which only AI can withstand. Tedious and repetitive tasks are delegated to the cybersecurity officers, as well. To change cybersecurity measures, AI needs to be the sole delegator in solving the problem of doing those menial tasks. It leads to the question of how capable will AI be in comparison to human delegators or other measures in place to fight cyber threats. The idea of enforcing AI comes to shape based on current security measures that are preventing cybercrimes and their effectiveness in stopping cybercrimes.

The first actant in this network is the Internet. The Internet, on its first basis, served as a communication tool and as an information provider. Now, it is the leading cause of increasing number of cybercrimes and variety in cybercrimes (Dilek, 2015). The Internet is open for hackers and malicious actors that prey on users in the Internet, to steal their information and misuse that information. The user are also part of this network, and the hackers are subparts of the users. These malicious actors are human, so they cannot understand a machine's language.

They have created human-readable programs that translates to machine languages, such as malware, virus, or worms, to penetrate the users on the Internet. There are countermeasures in place against such attacks, like anti-virus software. The anti-virus software detects and analyzes infected machines and prevents the infected program from running. However, Norton, an anti-virus provider, pointed out that Antivirus alone may not be enough. They state that the threat landscape is continuously changing, so approaches to protecting information must also change (Symanovich, n.d.). The next actants are cybersecurity officers, trained to detect and analyze malicious forms of attack. But there is a limit to what an officer can do against a highly sophisticated attack. As Kewlani (2018) states "cybersecurity officers have been barely successful in bringing down the time it takes to acknowledge an attack ‘after it has happened,' let alone stay ahead of the curve and predict the next breach or attack much before the attackers strike the first blow." The question remains, what could stop such evolving attacks?

The integration of Artificial Intelligence to point guard information is a potential solution to boost cybersecurity. A non-profit company called OpenAI is using deep learning techniques to train a bot to learn a game called Dota2 and it beat the world's top professional players who have been playing the game for more than three years. The game is very complex and takes an average player almost a year to learn. But the bots learned to play within ten months and beat the top players with ease (Verge, 2019). If the new cybersecurity embeds applications like Deep Learning, it could eliminate the task of repetitiveness and tediousness. As Bharadwaj (2019) points out, "identification and assessment of cyber threats require scouring through large volumes of data and looking for anomalous data points." It might take several days for humans to analyze and understand the data, while the AI can do it in much less time.

Many technologists say that human beings are the weakest link in security. According to a study done by IBM, human error is the cause of 49% of cybersecurity breaches (Cambridge, 2019). First, it is essential to distinguish the level of errors into two parts, skill-based and decision-based error. A skill-based error happens when a developer accidentally creates a bug. If the system is exploited from the bug's vulnerability, information within the system can be compromised. The prominent errors arise from decision-based failure, where users with their lack of skill in using a computer and lack of general knowledge make bad decisions. Instances, like creating a weak password that can be easily guessed, clicking links that are possible phishing sites, and sharing information with random websites (Ahola, 2019). So, it makes sense to build a security net that understands that users make mistakes and develop technologies that work around those issues. Such technology is the intelligent agent system. Intelligent agents obtain the user's input and try to understand the information to make a rational decision. The basis of error lies in the input the user provides. Still, if an intelligent agent system can learn what a correct input should be and requests the user to provide a correct input, it can eliminate user vulnerability (Rouse, n.d.). The password weakness detection tool is an example of an intelligent system. If a user tries to enter an easily guessable password, the system rejects that password until the user enters a much more complex password. Figure 2, shown below, illustrates the intelligent system implemented in Facebook to detect weak passwords. Bots are also considered intelligent agents, as they are already present in the current technologies like iPhone's Siri, Amazon's Alexa, and Google's Bixby. These devices have a smart understanding of the user's input and respond accordingly. Such an application can be quite useful to implement and can help mitigate risks caused by user inputs.

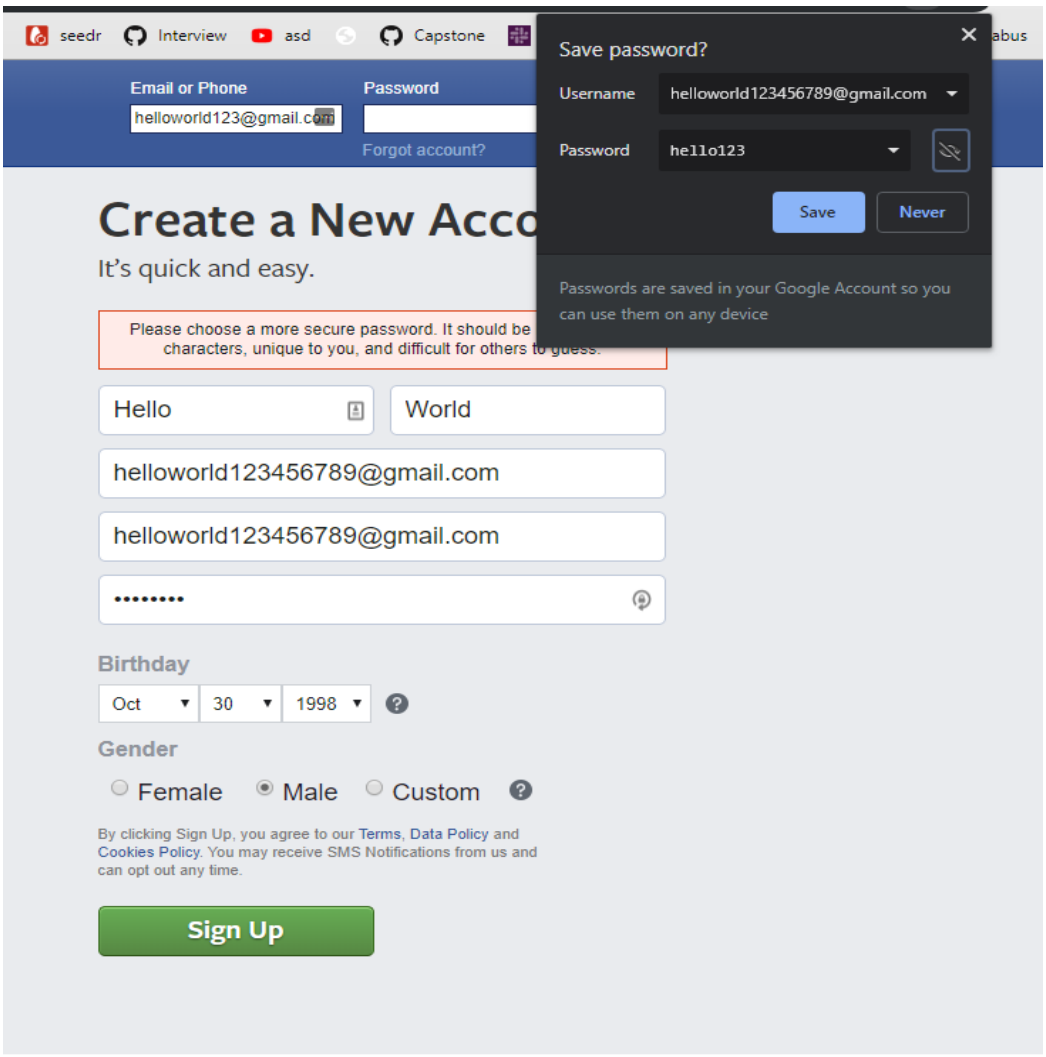

Figure 2. Entering a very insecure password in Facebook, blocked by an intelligent system (Parajuli, 2019)

While Artificial intelligence seems promising in developing a new form of cybersecurity, other nations are going even further by exploiting massive groups of people through the application of AI technologies. In the 2016 election, millions of Americans were exposed to fake news and propaganda. These influences affected the choice voters would make in picking a candidate for the 2016 election. To understand this attack observing the different forms of actants is also necessary. Twitter was the primary source of fake news and strategic propaganda. The attackers were identified to be the Russians, and their attacking method had different applications of AI in creating fake news and propaganda (Horowitz et al. 2018). Horowitz and

his group talk about various aspects of AI against these attacks. To counter disinformation and fake news, a new prospect is already making progress. In 2017, Google, Poynter, and MIT formed a partnership to research natural language processing and deep learning to detect misinformation. They built an algorithm to detect these nuances with an 80% success rate (The Poytner Institute, 2018).

Research question and methods

Cybersecurity faces new forms of threat every day, and the question of security is an everlasting problem as the information age progresses, so will the risks. These threats will not stagnate. They will evolve into more sophisticated attacks. A model must be drawn to analyze and understand the attacks. Through this model breaking down different parts of the attacks helps build a countermeasure, and AI is a promising new field to seek countermeasures. The question this research will address is How will Artificial Intelligence shape the landscape of cyber security and how effective will it be compared to current cyber security measures? This will provide more insight as to what will the security system look like and it will show a comparison between AI security measures and current security measures. It will also help measure the overall usefulness of the technology.

Reviewing case studies about the different applications of Artificial Intelligence will be used to identify what software methods is needed to address the variety of problems. Applications like search method, intrusion detection system and intelligent agent have shown the most promise and therefore they will be further studied. To understand more about the AI in cybersecurity, case laws will also be studied to see how the government will enforce laws in accordance to AI. This also includes the ethical questions that AI in security might raise.

Groups like Google's child company Chronicle, have started to research AI applications and their reports will be studied to gain knowledge about their tools that are in place to fight cyber threats (Chronicle, n.d.). Darktrace, another forefront company that is building AI security measures, have posted blogs about their technologies (Heinemeyer, 2019). The blogs will be studied to obtain more in depth understanding about their technologies. The blog's writers are experts in cyber security field and are working directly on the technologies. Understanding current technologies is also important, therefore studying the forefront cyber security technologies that are current will provide for a comparison with the AI technologies. While it is essential to see the usefulness of AI, it equally important to judge the effectiveness of these applications. Research done by Apruzzese et al. (2018) can show more insight to see if these applications are even useful in fighting cybercrimes. This research has pooled experiments to detect anomalies by using machine learning algorithms, and those observed data set could help understand the algorithms rate of success.

Conclusion

Cybersecurity is facing new challenges every day, and they face humans actants as well as non-human actants. Their ability to fight against threats is diminishing because of the increasing acceleration of attacks. Therefore, stronger mechanisms like Artificial Intelligence and their applications can prove to be useful. This research will provide more insight into applications of AI and the different forms of attack. It will also give a study of the attacks and an in-depth understanding of the actants, that are involved in combatting cyber threats as well as actants that draw out these attacks. The thesis will begin over December, first by obtaining information about the current cyber security and AI technologies. Then by January, case studies about AI

applications will be studied and their data will be observed to assess their effectiveness. In February, the comparison between the AI and current technologies will be conducted. Finally, the thesis paper will be completed by March or April.

References

Ahola, M. (2019). The Role of Human Error in Successful Cyber Security Breaches. *GetUsSecure*. Retrieved from https://blog.getusecure.com/post/the-role-of-human-error-in-successful-cyber-security-breaches

Apruzzese, G., Colajanni, M., Ferretti, L., Guido, A., & Marchetti, M. (2018). On the effectiveness of machine and deep learning for cyber security. *2018 10th International Conference on Cyber Conflict (CyCon)*. doi: 10.23919/cycon.2018.8405026

Bharadwaj, R. (2019). Artificial Intelligence in Cybersecurity – Current Use-Cases and Capabilities. *Emerj*. Retrieved from https://emerj.com/ai-sector-overviews/artificial-intelligence-cybersecurity/

Browser Market Share. (2016, May). *NetApplications*. Retrieved from https://netmarketshare.com/

Cambridge, Mass. (2019). IBM Study Shows Data Breach Costs on the Rise; Financial Impact Felt for Years. (2019). *IBM Newsroom*. Retrieved from https://newsroom.ibm.com/2019-07-23-IBM-Study-Shows-Data-Breach-Costs-on-the-Rise-Financial-Impact-Felt-for-Years

Chronicle. (n.d.). *X Development LLC*. Retrieved from https://x.company/projects/chronicle/

Dilek, S., Çakır, H., & Aydin, M. (2015, January). Applications of Artificial Intelligence Techniques to Combating Cyber Crimes: A Review. *Arixv*.

Farfan, B. (2019). The New "World's Largest Bookstore" Proves That Independents Beat Chains. Retrieved from https://www.thebalancesmb.com/is-barnes-amp-noble-the-worlds-largest-bookstore-2892133.

Harris, T. R. (2018). The Comprehensive Guide To College Textbook Trends [Infographic]. Retrieved from https://www.leadwinds.com/the-comprehensive-guide-to-college-textbook-trends/.

Horowitz, M. C., Allen, G. C., Saravalle, E., Cho, A., Fredrick, K., & Scharre, P. (2018). Artificial Intelligence and International Security. *CNAS*, volume(issue)1–27.

Kewlani, R. (2018). The evolving role of AI in effectively combating cybercrime. *Fractal*.

Kingkade, T. (2017). CHART: The INSANE Growth In College Textbook Prices. Retrieved from https://www.huffpost.com/entry/college-textbook-prices-increase_n_2409153.

Latour, B. (1992). Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts. In N. Name (Eds) *Shaping Technology/Building Society: Studies in Sociotechnical Change*, Cambridge, MA, MIT Press, pp. 151–180.

Heinemeyer, M. (2019). Stop the clock: How Autonomous Response contains cyber-threats in seconds. Retrieved from https://www.darktrace.com/en/blog/stop-the-clock-how-autonomous-response-contains-cyber-threats-in-seconds/.

Quick Facts (2017). *UVA Library*. Retrieved from https://www.library.virginia.edu/about-uva-library/.

Rouse, M., & Rouse, M. (2019) What is intelligent agent? - Definition from WhatIs.com. Retrieved from https://searchenterpriseai.techtarget.com/definition/agent-intelligent-agent.

Statt, N. (2019). OpenAI's Dota 2 AI steamrolls world champion e-sports team with back-to-back victories. *The Verge*.

Symanovich, S. (n.d.). Why antivirus may not be enough. *Symantec Corporation*. Retrieved from https://us.norton.com/internetsecurity-privacy-why-antivirus-may-not-be-enough.html

The Poynter Institute (2018). Poynter receives $3 million from Google to lead program teaching teens to tell fact from fiction online. *Poynter*. Retrieved from https://www.poynter.org/news-release/2018/poynter-receives-3-million-from-google-to-lead-program-teaching-teens-to-tell-fact-from-fiction-online/