

INFO1111: Computing 1A Professionalism

2023 Semester 2

Skills: Team Project Report

Submission number: 1

Github link: <https://github.com/Iambehindyou2/Info1111skill-new->

Team Members:

Name	Student ID	Levels already achieved	Levels being attempted	Selected Major
ZHANG, BOYUAN	530575630	X	A	Computer Science
LIU, YOUTAO	530038847	X	A	Data Science
HAO, CHANGYI	530162441	X	A	SW Development
ZHANG, KAILIN	530576006	X	A	Cyber Security

1. 1.1 Developing Industry Skills

- **1.YouTube tutorials** – There are many high quality YouTube tutorials nowadays from various creators. YouTube tutorials also extend from beginner level to advanced level, therefore it is applicable for almost all skill levels, when you are trying to learn a range of new skills. Compared to other approaches to learn a new skill, YouTube tutorials have no cost and are easily accessed as all you need to do is to type the interested skill into the search bar. We put YouTube tutorials at first place as it is effective and cheap.
- **2.Online courses** – There are many online courses nowadays that help students to learn the skills they wish to acquire. Those courses are charged, and generally have higher quality compared to YouTube Tutorials, as it has various teaching approaches such as online quizzes and tutors that help students along the way. However, such courses generally have a high cost, and therefore its quality and utility are generally exaggerated from advertisement, and courses generally cover an Important skill that consists of many other components, and many times those components are not that useful. Online courses are good, but not as good as YouTube Tutorials as it is not as focused, and it has a high cost.
- **3. Online websites** – When learning a new skill in tech, we usually search on google, and many times we come across high quality websites that serve educational purposes which are useful for us to learn a new skill. Such websites are mostly technical; therefore, it might not be as friendly as the previous 2 resources to beginners. Websites such like W3Schools teaches programming skills and there are also websites like Stackoverflow where you can discuss specific problems you meet during the course of learning. It is not as interesting to learn from websites compared to the previous 2 resources, but it is still a generally good.
- **4.ChatGPT** – This new online AI chatbot has a rapid growth in its popularity. It is useful resource to learn as you are asking all questions regarding to the skill you are trying to learn, and it will provide you with answers that are specific and generally correct. The problem with ChatGPT is that some information it provides is not always correct. People might become relied to it and lose the motivation to study for the new skill.
- **5.Books** – Due to the rapid development of technology, people tend to read less nowadays. However, it is still a good way to study as it provides the entire knowledge framework of specified topic. In the tech industry, there are discoveries everyday, and the problem with books is that it might not always be up to date. Books are not as efficient compared to other sources of learning, and the information it provides might be old and have no implementation in the current industry.

Contribution Overview Our group meets up every week to discuss our roles and help each other out. In the first few meetings, we discussed our strengths and weaknesses and allocated ourselves to roles that are suitable for us. After the first few meetings, it was clear what the individuals in the group had to do, and we just helped ourselves with problems we encountered when completing individuals' work. There were not any significant variations in the level of involvement, as all work was allocated equally so that each member of our group had the same workload as everyone else. For the first submission, we will just combine each group member's section into an entire document using LaTeX.

submission 2 contribution overview: In this submission, it was clear what each member is required to be as each of us work on individual tasks in corresponding fields. I first decided the tools to be discussed for level B, and we helped each other to update version control on github repository. We have our tutorial on Thursday and we go to the library after the tutorial every week to work on level B and help each other with the technical part.

2. Level A: Basic Skills

2.1. Skills for SW Development: HAO,CHANGYI

stronger skill

I think I am good at programming/software development, which is a skill on the SFIA list. This skill is a basic and key skill in the field of software development. The first reason is that if you master programming languages, development tools, and frameworks, it can increase the quality and efficiency of development, saving time and reducing mistakes when writing code. This can also increase the quantity of programming output. Software quality is very important because it is directly related to the reliability and security of a software system. The second reason is to better meet the needs of customers. For example, when customers approach our company for app development, having good skills in this area can help us easily understand their requirements and more easily satisfy their needs. To be honest, I am just slightly better at this than other people. This semester, I have a course to study Python, and I use my free time to study C++. In this case, I can use different languages in different situations. For instance, C++ is suitable for the development of underlying systems and high-performance computing programs like embedded systems, while Python is suitable for small-sized or original projects such as AI, automated scripts, and data science. In real-world conditions, we can use Python and C++ to write mixed code to maximize the quality of our code.

Weaker Skill

I think the main major defect is skill of Vulnerability research. The vulnerability research can help programmers find and fix the. This makes sure that the software you are developing is secure and is reliable and marketable. If you not mix this problems on time maybe some immoral hackers will attack the system that make the system paralyze. This leads some customers' personal information will be revealed. Therefore, the customers will prosecute the programmers who design these software. As a result the programmers who design these software give compensation to customers. In the worst case, if there are not fix bugs in the software, then legal liability will be pursued, because developers may have failed to take the necessary steps to protect user information. There are three ways to improve these skill. First of all I should study network security knowledge to improve this skill. There are many basic knowledge video about network security in YouTube. Therefore I learn the knowledge on line. Then I should learn the rudimentary knowledge of computer such as operating system and networking protocol. When we learn about the rudimentary knowledge of computer we can understand why these bugs will be produced. Moreover how to use the programming to fix different kinds of bugs. The third things I need do is research vulnerabilities. Because when I learn this knowledge it can prevent the same bug from happening again.

2.2. Skills for CyberSecurity: ZHANG, KAILIN

stronger skill

I chose Threat Intelligence from the SFIA skills list because I think it is relevant to my major, Cybersecurity. Cybersecurity's main goal in general is to protect the user's computer and network from attack, and detecting potential threats is a key element in preventing the attack. It is essential for people in Cybersecurity to learn threat intelligence, it covers gathering data, processing threat data and enabling security[1]. These elements are without a doubt relevant to a cybersecurity major. Frankly speaking I don't have a

strength that can let me stand out in threat intelligence yet but I do have some advantages in learning threat intelligence. First of all, I am learning R in one of my data science courses. R is a programming language for statistical computing. In my own words, R is very helpful in manipulating, processing and packaging data for data analysis, etc. A big part of cybersecurity is to protect users' data and information, this also involves threat intelligence. By sufficiently learning R, I will be able to deal with the data gathered from various sources and by having a deeper learning with data I will be able to know how to protect the data as well. Secondly, I chose PHP for my self-learning topic, which is a general-purpose scripting language for building a web. It is very common for people nowadays to use the internet and surf the web, which might contain potential threats embedded in different websites. Knowing how the PHP coding works can help me further detect the potential threats and enable security tools for prevention.

Weaker Skill

The skill that I feel like I am currently weak in but still plays a major part in Cybersecurity is Digital Forensics. Digital forensics involves finding evidence on computers and other devices[2]. Evidence that is used to support one's illegal act by using the internet, which is also what Cybersecurity people do for their job. Preventing one's data and information leakage, and if someone attacks and tries to get information and data illegally Cyber securities will then try to trace where is the information leaked and gather data for evidence supporting this "criminal" act. In digital Forensics, data collection is also a main part. Trying to dig up the small evidence requires the user to gather data from the usage of computers and the internet, which needs users to have a good understanding in handling data. One of the things that I can improve is the way of collecting data. Web scraping and web crawling is a required skill in my opinion for Digital Forensic that I have to know and learn how to use. Web scraping is extracting data from websites and web crawling is finding and discovering url[4]. These two techniques will be efficient in collecting data from the internet. Next, I have to improve my scripting and coding language, for example Javascript, python and other languages. For example, web scraping requires a solid knowledge of python, so learning different relevant coding language will improve my capability in digital forensics.

2.3. Skills for Computer Science: ZHANG,BOYUAN

stronger skill

The skill that I believe I am currently the strongest is numerical analysis. Numerical analysis is necessary for the major of computer science, because in areas such as machine learning, it requires to train computers to learn from data, which heavily relies on mathematical concepts such as linear algebra, sets and probability. Mathematics has been one of my strongest subjects in high-school and first year of university, as I achieved great grades in mathematics related subjects, therefore I am comfortable with numerical analysis. I am currently taking COMP2123, which is a course about data structures and algorithms, we were introduced the concepts of iterations and runtime, and I was able to understand that area of knowledge in a fast pace. Such concepts are crucial in the major of computer science, as it will help with the development of efficient algorithms, and apply numerical optimization, which is critical in many areas of computing, such as robotics and machine learning. Overall, numerical analysis is important in many areas around computer sciences, especially algorithm design and machine learning.

Weaker Skill

The skill in which I believe I am currently the weakest is innovation. Innovations leads to the creation of new technologies in which we can use on computers, such as artificial intelligence. The tech industry is also based on innovations from new ideas and product; therefore, innovation is crucial in the major of computer science. I usually find myself good at following instructions, and lack of creativity. To improve my innovation, I think I will try communicating more with classmates, so that we can discuss each other's ideas and perspective, so that my thinking will become various as I can think from many other perspectives, which should improve my innovation. I should also start taking more risks and learn from failures so that I can get more insights on decision making and think from different perspectives. I can also start daydreaming, although it may sound ridiculous, but in this rapidly changing society, anything can possibly happen with the help of technology. Overall, innovation plays in a crucial role in the computer science major, and it is also an area I need to improve on, I will try to get into more discussions and dream big.

2.4. Skills for Data Science: LIU, YOUTAO

stronger skill

Firstly, Data visualisation enables people to clearly understand concepts, ideas, and facts through charts and photograph. Obviously, R studio is a great tool for converting data into charts. In R studio, we can use some codes to upload the database and plot a graph, such as ggplot and barplot. The reason why data visualisation is significant is that it can clearly demonstrate the correlation between data and easily identify trends and outliers of data. Besides, the data visualisation can help the public to comprehend the data, compared to numbers or text, charts are easier for people to remember and understand. For instance, a college student who wants to start a business can make decisions by data visualization to find the sales volume of different products in a certain industry and the sales situation in different seasons and regions. Currently, I think this parts is what I am best at, the R studio does not require much programming foundation, but it can easily upload database into charts using codes. For example, if I want to figure out some questions, such as how much money they spend on their food a day, I have to make a survey to get the data,

and then I collect the data and upload it to R studio, then I can get graphs by using R language.

Weaker Skill

Data modelling is a basic part of data science, because it can help people to comprehend the correlation between varieties and analyse the trend. Besides, the data modelling is very important to develop predictive modelling, we can analyze future trends and make right decisions by using the predictive modelling. We can find potential profit or risk by predicting the future results. Then I figure out that the data modelling is very crucial for the data science, which includes the data analyzing, developing predictive modelling and improve the data management strategies. However, I think this part is currently weakest for me, so I can learn it by watching the videos that are useful for the data modelling, besides I can learn the data modelling on the website, such as Data camp, Coursera, and Chatgpt, these website helps a lot to me. Personally, I would say that the data science covers so many knowledge so that every part seems to be very crucial, however, I the data visualization and data modelling is the most significant to the data science in my view.

3. Level B: Tools

3.1. Tools for SD:CHANGYIHAO

The first tool I want use is the IDE. The main functionality of the IDE is use to integrate many different tools in a same place. We can use IDE writing programming, debug programming and running coding. For example, the IDE have editor it can writing code, the IDE have compiler it can compiles code, Also we can use the IDE to debug coding. The programmer can finish these steps in the IDE. According to the IDE provided a integrated development program, we can use IDE in many different situations in software developing. like the Desktop application development, mobile application development, and so on.² The main of weakness of IDE is the hard to run if we not use better or higher system sources [1] Because IDE merge many different tools therefor, its memory and processor usage can be high. As a result, it may lead the computer slow down or crashes. Also the limitation of the IDE is that each IDE only depend on it special language. For the instance , a IDE use for C++ may not use for the Java. [2]

The second tool is the VCS(version control system).The main functionality of the VCS is tracking and management the history of different versions of coding. Also it can let lots of people writing a same code together and if the programmer have demand it can back to previous versions [3] The ways of VCS used: The software developers can use VCS to manage the version history of code without overwriting the versions. And the programmer can use different branches and tags to control different step of you developing and release time. In this case we can multiple versions at the same time The weakness of the VCS is it may happen mistake or conflict when we merge it. Therefore, we should solve this problem by ourselves Also the security of VCS code base is important. When use the VCS we should take some steps to protect like limit other people face the code base of access [4]

3.2. Tools for CyberSecurity: ZHANG, KAILIN

The first tool is kali linux, Kali Linux is a Debian-derived Linux distribution designed for digital forensics and penetration testing. Kali Linux is more like a platform rather than tools and operating system. Kali Linux is widely used in the cybersecurity field, especially for cyber experts and ethical hackers[5]. Since Kali Linux is open-source and already has more than 600 penetration testing tools pre installed, people can easily gain access to the Kali Linux and use it to test security controls in systems and networks. Kali Linux can also analyse digital evidence, the behaviour of malware or other software and investigate cyber incidents. In addition, due to its wide range of cybersecurity tools pre installed, people can use it to crack the password, scan for vulnerability and also network sniff[6]. The biggest disadvantage for Kali Linux is because of its system. Kali Linux is not ideal for general-purpose computing because it is specifically made for security professions and it is not friendly to those who are new to the Linux system. Plus Kali Linux itself is slow and may contain software malfunctions[7].

The second tool is called John the Ripper, it is a widely-used password cracking tool in the cybersecurity industry. John the Ripper is an open-source tool that is currently available on many operating systems.Its main purpose is to detect weak Unix passwords. One way of using John the Ripper is for password security auditing. By obtaining a list of hashed passwords from a system, involving extracting the passwords from the database. The second way of using it is for password recovery, identifying the file containing the lost password and running John the Ripper to recover the password[8]. The limitation for John the Ripper is that it will be time-consuming and not efficient especially towards

those complex passwords. Next the entire program depends on the processing power of the hardware, which is a limitation. Last, the program may involve ethical issues and without permission may cause illegal consequences[9].

3.3. Tool for Data Science: LIU, YOUTAO

I would like to introduce R studio first. The main function of the R studio is that R studio provides a lot of R libraries and tools which support data processing and visualization, and more. Besides, R studio can support create interactive and dynamic data visualization. R studio can be used in many areas, such as math, statistics, and data modelling. We can use R studio to create a graph by using the codes “ggplot”, then it can show the data very clearly[10]. However, there are some limitations. R studio has some barriers in coding, which means the users have to optimize their codes to get higher efficiency, and R studio may be difficult to those people who are not familiar with statistical computing[11].

For Tableau, it is a very useful tool to data science. It can help users create interactive and dynamic dashboards and reports. Tableau provides a series of visualization tools which can help users analyze the data more clearly, besides Tableau support different types of data, like SQL and Excel. Mostly, Tableau are used in financial area and medical area, for instance, some people in the stock market may use Tableau to better explore and analyze stock market data[12]. Tableau provides rich visualization tools that can help users better understand stock market data such as stock prices, trading volume, and market trends. Through Tableau, users can import stock market data from different data sources and create interactive dashboards and reports using its visualization tools to analyze the dynamics and trends of the stock market, helping them make more informed trading decisions. However, stock market analysis requires deeper data analysis and professional knowledge, and Tableau is just one of the tools used in this field. However, the Tableau may be expensive for some small company, and it is not suitable for data-analyzing in a huge amount[13].

3.4. Tools for Computer Science: ZHANG,BOYUAN

The first tool in the field of computer science is visual studio code (VSC). Visual studio code is a cross platform and open-sourced code editor developed by Microsoft. Its main functionalities include coding, debugging, and building applications from many platforms[14]. To start code, the users need to enter the application and select the programming language and download its corresponding environment and debugging extensions. With extensions, visual studio code can highlight, auto complete and code snippets to make programming more efficient with the given language syntax. Visual studio code also has a built in debugger that is used to identify errors in users’ code, it also provides the functionality of variables, break points to go through the program line by line. Visual studio code can also manage projects such as creating and managing folders, files, and work with others. While visual studio code is one of the most popular code editors, it still has many weaknesses. Memory usage is a weakness as visual studio code can be memory intensive, when running big projects can lead to crashes or low performance.

The second tool I chose is git. Git is a version control system that allows developers to share codebase and work with each other. Its main functionality is version control and collaboration. Git allows users to keep different versions and track changes to code. After installing git, the user needs to initialize a git repository using the git init command[15]. After repository is initialized, users can now stage and commit changes using git add command and commit the changes to the repository using the git commit

command. Git allows users to collaborate with others, users can share repositories with others by pushing it into GitHub, other contributors can clone the repository to local and make changes. Git is a widely used version control system, however it still has many weaknesses. Git has the problem of merge conflicts, when multiple users work on the same codebase, merge conflict can be a commonly occurring issue. Git also has a steep learning curve, so that for new users who are beginners, the functionalities can be challenging and difficult to use[16].

article

Bibliography

- [1] Salfarisi, “"advantage and disadvantage of using ide",” 2010, "See <https://salfarisi25.wordpress.com/2010/12/22/advantage-and-disadvantage-of-using-ide/>".
- [2] Codecademy, “"what is an ide?,"” 2023.
- [3] Atlassian, “"what is version control?,"” 2023.
- [4] Brightest, “"advantages of version control systems",” 2023.
- [5] O. Security, “Kali linux,” 2023, see <https://www.kali.org/>.
- [6] javatpoint, “Advantages and disadvantages of kali linux,” 2023.
- [7] TechQuintal, “Advantages and disadvantages of kali linux,” "See <https://www.techquintal.com/advantages-and-disadvantages-of-kali-linux/>".
- [8] GitHub, “John the ripper password cracker,” see <https://github.com/openwall/john>.
- [9] S. Weakness, “Hacking tools: John the ripper,” see <https://systemweakness.com/hacking-tools-john-the-ripper-806a196e3af5/>.
- [10] Zhouyunlaijiushiwo, “Introduction to rstudio,” 2020, see <https://www.jianshu.com/p/b8b4b44bb352>.
- [11] Bkash, “A version of this package for your version of r might be available,” 2023, see <https://community.rstudio.com/t/a-version-of-this-package-for-your-version-of-r-might-be-available-elsewhere/166309>.
- [12] Tableau, “Getting started with tableau desktop,” see <https://help.tableau.com/current/guides/get-started-tutorial/zh-cn/get-started-tutorial-home.htm>.
- [13] vgherard, “Model misspecification and linear sandwiches,” 2023, see <https://www.r-bloggers.com/>.
- [14] T. Team, “Is visual studio code really the best code editor?” 2021, see <https://www.tabnine.com/blog/visual-studio-code-really-the-best-code-editor/>.
- [15] Atlassian, “What is git,” 2023, see <https://www.atlassian.com/git/tutorials/what-is-git>.
- [16] —, “Git merge conflicts,” 2023, see <https://www.atlassian.com/git/tutorials/using-branches/merge-conflicts>.