ISSN: 2071-2987
(Online version)
2223-523X
(Print version)

IJDATICS

International Journal of Design, Analysis and Tools for Integrated Circuits and Systems

INTERNATIONAL JOURNAL OF DESIGN, ANALYSIS AND TOOLS FOR INTEGRATED CIRCUITS AND SYSTEMS

The International Journal of Design, Analysis and Tools for Integrated Circuits and Systems (IJDATICS) was created by a network of researchers and engineers both from academia and industry. IJDATICS is an international journal intended for professionals and researchers in all fields of design, analysis and tools for integrated circuits and systems. The objective of the IJDATICS is to serve a better understanding between the community of researchers and practitioners both from academia and industry.

Publisher

Cooperation Name	: Solari Co., Hong Kong
Address	: Unit 1-5, 20/F, Midas Plaza, 1 Tai Yau Street, San Po Kong, Kowloon, Hong Kong
Phone	: (852) 3966-2536

Preface

Welcome to the Volume 13 Number 1 of the International Journal of Design, Analysis and Tools for Integrated Circuits and Systems (IJDATICS). This issue presents six high quality academic papers, providing a well-rounded snapshot of current research in the field of Computing in AI, Internet of Things (IoT), Integrated Circuits and Systems and Computer Engineering Technology.

There are two key themes evident in these papers:

- ***Computer security:*** Two papers investigate how artificial intelligence can be used to ensure the computer and user security.
- ***AI Application:*** Five articles focusing on practical applications of computer and artificial intelligence technologies.

We would also like to thank the IJDATICS editorial team, which is led by:

Table of Contents

Vol.13,No.1,February 2024

Computer Forensics of Hacked Websites

Yu Qiao and Gabriela Mogos

***Abstract*— With the development of technology, more and more families are connected into a big digital world through the Internet which brings about an increasing number of targets for hackers to fulfill crime. Almost everyone in each family lives rely on the latest network applications such as e-bank, e-library, e-mail and so on. The more dependent on the web applications the more likely we are being attacked on the Internet. Not only the users of the web applications and website will face cyberattack but also the companies who carried out these applications and website meet hackers' threats. Hacking attempts are a common illegal action recently because most hackers hack into people's computer or companies' database for not only extort a respectable sum of money but also some classified documents. Internet hacking is happening everyday all over the world except a computer can be locked in a finite place with limited access and almost not connected to the outside Internet world.**

In this paper, we use attack scenarios on a vulnerable web application with widespread types of attacks: SQL Injection attack and XSS attack which are usually performed by using a web browser. We tried to set up some testing target machines such as DWVA and OWASP and establish a simple victim website and a simple detecting code in the end. Also, we use computer forensic analysis technic of attacker and victim machine to find some clues and finally try to find some possible ways to avoid websites being attacked. Further details will be mentioned in the methodology and result part and finally a future work section will be carried out to conclude the insufficient of this project and some future jobs.

***Index Terms*— Computer crime, Digital forensics, SQL injection, Database, Web-application attacks, XSS attacks.**

I. INTRODUCTION

If a company of web application wants to remit a case to the Court of Appeal due to accorporate espionage which lead to some important images and contracts' detail information and database are falsely obtained by hackers. Depending on the investigation, they must understand and apply a large number of legal concepts and precedents, such as chains of custody, destruction of evidence, and handling the presentation of evidence in court [1]. Digital evidence is an indispensable tool in order to collect more evidence of hacking. Digital evidence using a process which includes identifying, preserving, examining and analyzing. It has been scientifically validated and validated, and ultimately presented in court to answer certain legal questions.

Nowadays, there are so many hacking methods testing the database's security for instance, Cross-site Scripting, SQL injection, Code injection, Buffer Overflow [6]. The following table I shows some basic web application attack statements and their attacking types.

All authors are with the Department of Computing, School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, China. (email: Gabriela.Mogos@xjtlu.edu.cn).

Table I. List of common web application attack statements

Attack statements	Attack type
redirect : %25{ (new+java.lang.ProcessBuilder (new+java.lang.String[] {'command', 'goes', 'here'})) .start ( ) }	struts2 Arbitrary command execute
<img src=javascript: alert ('xss') >	Cross-site Scripting
'><script>alert (document.cookie) </script>	Cross-site Scripting
" or "a"="a	SQL injection
1' and 1=2 union select * from	SQL injection

Attackers using similar statements invade and destroy not only website but also other web application's database which causes the leakage of users' information and even Loss of property.

The SQLi attack is a kind of attack which poses a serious security threat to every web application and website through the database and takes control of the database server of the web application. It leverages code to leverage the Web site by changing back-end SQL statements in the application database layer [7].

According to the report of a power enterprise, almost no dispatching data network applications could escape from suffering different forms of cyberattacks, of which more than 60% are SQL injection attacks and the number of SQL injection attacks on these network applications is still rising year by year [8].

In addition, according to the OWASP's "Top Ten security vulnerability list", which made a list of the most critical Web application vulnerabilities, since 2013, not only session management and cross-site scripting are on the list, but also the injection type attack especially the SQL injection [9,10].

SQL injection is utilizing code for attackers inserts malicious information into a user's input parameters. Then attackers can input or output the data or even update and delete some data in the database.

There are two types of SQL injection: integer and string, integer type Indicates that the parameter entered by the user is an integer and don't have single quotes in statements. String type Indicates that the parameter entered by the user is a string with corresponding symbol such as Double quotation marks, parentheses added to the statement.

The following table II shows some other types of SQL injection and a sample attack statement [7].

Type	Description	Techniques
Tautology	Use a conditional statement which is always true.	SELECT * FROM table WHERE login= '' or 1=1--
Logically Illegal Incorrect	Inject codes in vulnerable or injectable parameters which creates syntax, type conversion, or logical error.	SELECT * FROM table WHERE login 'kao''' AND password =
Union	Combine the result sets of two or more statements	SELECT * FROM table WHERE login='' UNION SELECT ** FROM table WHERE No=12500 -- AND password ='' AND pin=
Piggy-Backed	Insert additional queries to the original statement	SELECT * FROM table WHERE login—'kao' AND password=''; drop table users -- ' AND pin=223
Stored Procedure	Execute built-in functions using malicious SQL codes	SELECT * FROM table WHERE login= 'kao' AND password =gal·; SHUTDOWN;--;
Alternate Encodings	Modify the injection statement by alternating encoding to escape from detection	SELECT * FROM table WHERE login= ' kw' ;exec(c har(0x73697 574646f776e)) —` AND password -='lai' AND pin =; SHUTDOWN;--;

Table II. types of SQLi attack [7]

Furthermore, the following figure 1 shows the how a SQL injection statement works [6]. Firstly, an attacker sends the harmful HTTP input to the web. The web application server then constructs the SQL statement and sends it to the database.

After that, the database gets the SQL statement and uses some basic SQL statement such as 'SELECT * FROM table WHERE login = `kao' AND password' in order to response to the web application server. The web application server returns some information of the database to the aggressor so that the attackers finally gain the data.

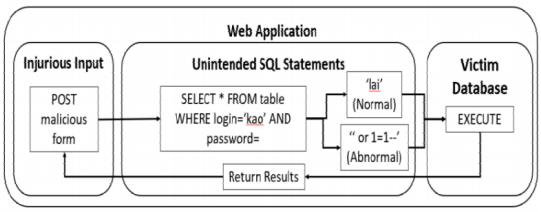


Figure 1. Procedures of SQLi statements [6]

XSS attack is another kind of attacking method with high risk which can not only get the web applications' database but also inject HTML scripts and cookies.

XSS attack [12] can tamper with the website and get users' database through cookie with JavaScript.

II. METHODOLOGY

To simulate a scene of website hacking and web application forensics, we first tried a SQL injection on the OWASP and DWVA platform. The IP address of the attack machine is 192.168.56.102. And the Victim IP address of OWASP is: 192.168.3.33. In this project, the process of SQL injection is mentioned in figure 1. First to determines whether the current web page can be SQL injected using some basic statements such as add the statement "and 1=2 union select 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15," or "and 1=2 UNION SELECT 1,2,3,4,5,6,7,8,9, *group_concat* (table_name), 11,12,13,14,15,16 from information schema.tables where table schema=0x3428296E6A"after the URL bar.

If this webpage can be easily injected, the website will response with little change. Then we will find the SQL injection point by utilizing the SqlMap which is an automatic injection tool that can scan, find out, and inject the given URL with SQLi loophole. After finding the inject point, we can easily continue determine the database type and database mode due to whether the type of the SQL injection is integer type or string type.

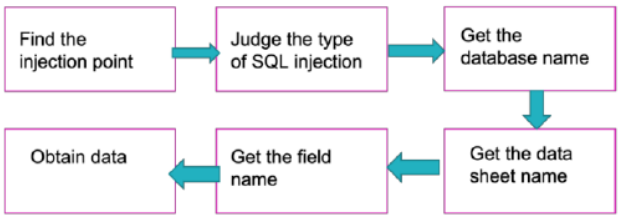


Fig. 2. Flowchart of SQL Injection penetration.

As if the SQLi is a string type, we could try to import "3' or 1=1#". If the webpage makes a response, this means that the original argument is stored in a closed pair of single quotes, the single quote on the right is commented by "#", and the single quote on the left is closed by the number 2 and the statement is executed correctly after the input argument is closed. Thus, the inject point is a string type.

Moreover, we can get the database library name, table name, list name, and data with SQL statement after getting the type of the injection. While sometimes the database may be encrypted with algorithms. For instance, the MD5 algorithm will lead to the result of the data we get a long paragraph of numbers mixed with letters. Thus, we have to decode the data we got in the final.

After utilizing the OWASP and DWVA platform, a sample website is in used for a SQLi attack. For trying a SQL injection attack in a website, we used a raspberry pi machine as an attacker machine with Kali Linux system and another victim

machine build the sample website and server which has an IP address of 47.102.113.37.

The sample website is built with MySql database and some other plugins with known leakage which is out of date so that it ensures the injection could be easily happened.

The process of the attacking is similar to the process enforced on OWASP, which is:

1. Determines whether the current web page can be SQL injected;
2. Find the SQL injection point;
3. Determine the database type and database mode;
4. Get the database library name, table name, list name, and data;
5. Decode the data.

For this time, we tried to use SQLi attack on the login pages of the sample website. The first-time injection on the login form is done by injecting a known string "select * from tl where id = '1' and '1'='1'", which allow us to login this website directly without input a concerted password as a common user.

Figure 3 shows the database table and column name of the sample website we got from after implement SQL injection with sqlmap using statement: sqlmap -u http://47.102.113.37/ login.php?id=1 -D db -T.

For the XSS attack simulation, we utilize the same website with a similar beginning injection method with SQL injection. "select * from tl where id = '1' and '1'='1'" is used for the second time for further operation. We use remote file inclusion using the browser and upload two public accessible shells b374k-shell.php We can run the shells using admin page and null byte injection, for example http://47.102.113.37/admin/index.php? page=http:// 47.102.113.37/upload/b374/b374% 00.php. By utilizing this shell, we have removed the folder testl (nn -rf testl) and we have created new folder htest2.

After implementing hacking simulation on the websites, forensics is the following important stage. In forensics part, collecting information, checking, analyzing victim and attacker are basic steps. There are some methods to forensics from the log such as System log analysis, MySql log analysis, web shell extraction, grep.

The figure 3 shows some results from log analysis and interne history files from attacker's machine. These results are just extract from the attack machine for the log and to simulate that we find a real thread from the possible hackers' machine. In this case, several URL records are found, which proves that the attacked site was being visited.

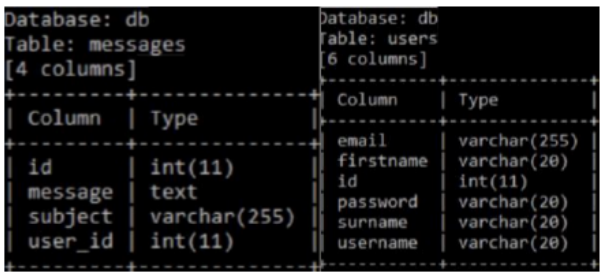


Fig. 3. Table and column name in SqlMap.

III. RESULTS

Although computer forensics contributes to decrease the possibility of websites from being attacked, we still need to realize our web applications are hacked. Thus, a simple code is written for detecting the SQL injection and XSS.

The part of the code is in the appendix part and some working methods will be mentioned in the following part. This code can detect attacks through detecting and filtering users' parameter. Parameter name and value of *getParameter* method and *getHeader* method are covered by XSS filtration. In addition, the half angle symbol which may easily lead to an injection weakness is replaced by full angle symbol.

Based on the research, we recommended several steps that the website could take to improve its security posture and prevent not only SQLi attacks but also other digital attacks in the future.

1. Strictly filter the input parameters or validate the input data and precompile using SQL statements [9]. Because most SQL injection attacks begin with inputting strings or other SQL statements on the URL part or other importable part. Thus, we can limit users only input numbers instead of strings to decrease being injected. In addition, the precompile methods ensure that the syntax structure of the SQL statement is not changed. Even if the user constructs the concatenated malicious statement, it will only be treated as a string of characters and will not be executed as an SQL command.

2. Regularly Update Software: Actually, the sample website is built with some other plugins which are out of date so that it ensures the injection could be easily happened. Many hacks occur because attackers can exploit known vulnerabilities in software that has not been updated.

To address this, we recommended that the website implement a regular software update program. This program should include regular checks for updates and patches, and a schedule for installing them.

3. Conduct Regular Security Audits: In order to maintain a high level of security, it is important to conduct regular security audits. These audits should include both internal and external assessments of the website's security posture. This will help to identify vulnerabilities and ensure that security controls are working effectively.

4. Engage the Services of a Third-Party Security Firm: Finally, we recommended that the website engage the services of a third-party security firm to provide ongoing security monitoring and guidance. This firm should have expertise in the latest security trends and best practices and should be able to provide regular reports on the website's security posture. They should also be available to provide guidance on how to address any vulnerabilities that are identified.

By implementing these recommendations, the website can significantly improve its security posture and reduce the risk of future hacks. It is important to note that security is an ongoing process, and it is important to continuously monitor and update security controls to address new threats as they emerge.

IV. CONCLUSIONS

Overall, hacking of the website really disturbs security problems to the maintainer and user of Web application.

This paper presents the harm of website hacking and why we should do hacking forensics after hacking happened. Moreover, the network penetration testing can be thorough from both attack and defensive aspects to evaluate the security status of the target website for the target and comprehensive security protection and reinforcement suggestions.

Web application forensics is a branch of digital forensics [6]. Forensic web application security attacks could be carried out more easily with a correct approach. However, there is still something that needs to be fulfilled in this project in the future work in order to have a comprehensive cognition of website forensics.

REFERENCES

[1] A. Philipp, D. Cowen, C.C. Davis, *Hacking exposed computer forensics*, 2nd ed. McGraw-Hill. 2010.

[2] A. Roberto. Hackers Cybercrime - Computer Security: Ethical Hacking, *ARIS2 - Advanced Research on Information Systems Security*, vol. 1, no. 1, pp.50-61, 2021.

[3] A.R. Caesarano, and I. Riadi, Network Forensics for Detecting SQL Injection Attacks Using NIST Method, *International Journal of Cyber-Security and Digital Forensics*, 7(4), 436+, 2018. available: https://link.gale.com/apps/doc/A603050347/AONE?u=anon-99f3Oacb&sid=googleScholar&xid=07ab3f8f [accessed 24 Apr 2023].

[4] D. Mualfah, I. Riadi, I. Network Forensics for Detecting Flooding Attack on Web Server, *International Journal of Computer Science and Information Security*. 15(2), 326-332, 2017. [13]

[5] M. Z. Gunduz, Biliim suclanna yonelik IP tabanh delil tespiti- IP-based Evidence Detection, *Master Dissertation, University of First,* 2013.

[6] T. Lokesh Sai Reddy, Forensics on a Hacked Website *International Journal of Innovations in Engineering Research and Technology*, vol. 8, No. 09, pp. 11-15, 2021.

[7] D.-Y. Kao, C.-J. Lai and C.-W Su. A Framework for SQL Injection Investigations: Detection, Investigation, and Forensics, *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2018. doi:10.1109/SMC.2018.00483.

[8] J. Sheng, Research on SQL Injection Attack and Defense Technology of Power Dispatching Data Network: Based on Data Mining, *Mobile Information Systems*, 2022. doi:10.1155/2022/6207275.

[9] JuHua.Y and Hong.Q (2022) 'SQL Injection penetration test based on DVWA platform', Science and Technology & Innovation, (21), pp. 71 - 73. doi:10.15913/j.cnki.kjycx.2022.21.022.

[10] N. Suteva, A. Mileva and M. Loleski, Computer forensic analisys of some web attacks, *World Congress on Internet Security (WorldCIS-2014), Internet Security (WorldCIS)*, 2014. doi:10.1109/WorldCIS.2014.7028164.

[11] A. Varol and Y.O. Sonmez, Review of Evidence Collection and Protection Phases in Digital Forensics Process, *International Journal of Information Security Science*, 6(4), pp. 39-46, 2017. Available at: https://search.ebscohost.com/login.aspx?direct=true&db=asn&AN=127673730&site=eds-live&scope=site [accessed: 25 April 2023].

[12] B. Gogoi, T. Ahmed, and H.K. Saikia, Detection of XSS attacks in web applications: a machine learning approach, *International Journal of Innovative Research in Computer Science & Technology*, vol. 9, no. 1, pp. 1-10, 2021, doi:10.21276/ijircst.2021.9.1.1.

Identity Reinvented: An Examination of Self-Sovereign Identity Using Blockchain for Improved User Privacy and Security

Aliyu Abubakar

***Abstract*— In today's digitally driven world, where data is often the currency of the online realm, safeguarding personal information has become paramount. The traditional models of identity management, centralized and prone to vulnerabilities, are now facing a formidable challenger - Self-Sovereign Identity (SSI). This paper embarks on a comprehensive journey through the realms of SSI, shedding light on its principles, the role of blockchain technology, real-world use cases, privacy and security enhancements, challenges and solutions, comparative analysis of blockchain platforms, illuminating case studies, future prospects, and the evolving regulatory landscape. By the end of this exploration, the reader will gain a profound understanding of how SSI on blockchain is poised to revolutionize digital identity management.**

***Index Terms*— blockchain, web3, security, identity, user privacy, SSI.**

I. INTRODUCTION

In the 21st century, cybersecurity has become a critical issue impacting individuals, organizations, and nations worldwide. As digital technologies continue advancing rapidly, so too do the threats lurking in the shadows of cyberspace. From massive data breaches to crippling ransomware attacks, cyber incidents can have devastating consequences. Understanding the evolving threat landscape is essential for developing effective defences. This paper provides a comprehensive analysis of current and emerging cybersecurity threats, highlighting major trends and their implications. The examination covers various threat types, attack motivations, technological advances in security, data privacy concerns, infrastructure vulnerabilities, human factors, regulatory frameworks, and future directions. By navigating the shadows of the cyber world, we can shed light on risks and work collectively toward a more secure digital future.

II. EMERGING THREAT LANDSCAPE

The cybersecurity threat landscape encompasses the range of potential attacks and vulnerabilities that exist within cyberspace. This landscape has transformed dramatically from isolated hackers in the 1990s to today's highly sophisticated cybercriminal networks. Major evolutionary milestones include the Morris Worm incident in 1988, rapid growth of the internet and e-commerce in the late 1990s, escalating politically motivated attacks in the 2000s, and the rise of ransomware and supply chain attacks in the 2010s [1]. Recent years have seen threat actors become more organized, strategic, and destructive.

Some high-profile cyberattacks in the 21st century include:

• The 2016 DNS breach at Dyn, which disrupted access to major websites including Twitter, Netflix, CNN, and others. This DDoS attack leveraged an insecure IoT botnet [2].

• The 2017 WannaCry and NotPetya ransomware outbreaks which affected hundreds of thousands of systems globally. WannaCry encrypted data and demanded ransom payments in bitcoin, while NotPetya caused over $10 billion in damages [3].

• The 2018 Marriott data breach exposed over 500 million guest records, making it one of the largest known breaches impacting an enterprise [4].

• The 2020 FireEye breach resulted in cybercriminals stealing sensitive tools used by governments to test their cyber defences. This supply chain attack highlighted exposures in the security industry [5].

Attack categories span malware, phishing, denial-of-service (DoS), man-in-the-middle attacks, and more. Threat actors range from cybercriminals seeking financial gain to hacktivists and nation-states engaging in espionage or disruption of critical infrastructure. With billions lost annually to cybercrime, the landscape remains fraught with risk [6].

III. TECHNOLOGICAL ADVANCES IN CYBERSECURITY

As cyber threats have advanced, so too have cyber defences through applied research and innovation. Security technologies provide vital visibility, while enabling rapid detection and response to mitigate impacts of attacks.

Artificial intelligence (AI) and machine learning systems can identify threats missed by traditional signature-based tools. For instance, Darktrace's AI models normal network activity to pinpoint anomalies indicative of emerging cyber risks [7]. Cloud-based threat intelligence platforms like Recorded Future asset real-time monitoring of threat actor communications on the dark web to provide actionable intelligence [8].

Data protection relies increasingly on encryption, secure multi-party computation (MPC), and blockchain technologies to safeguard confidentiality and integrity. MPC enables collaborative analytics on encrypted data without exposing raw data. Blockchains establish decentralized trust mechanisms, as seen in cybersecurity use cases like Keyless SSL and Certificate Transparency monitoring [9].

Applied effectively, modern security technologies provide significant advantages against evolving threats. However, people, processes and partnerships remain essential elements.

IV. DATA PRIVACY AND COMPLIANCE

With expanding digital footprints, escalating data breaches, and increased reliance on data analytics, concerns around data privacy have heightened globally. Regulations like the EU's General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA) aim to enhance data protections and user privacy rights.

For organizations, maintaining compliance with data regulations poses numerous challenges, from data discovery and legacy systems to vendor assessments and access controls [10]. Non-compliance risks hefty financial penalties, along with reputation loss and customer churn.

Encryption provides a safeguard by securing data in transit and at rest. With robust key management, encrypted data stays protected even when breached or improperly accessed. Firms like CloudKnox and CyberArk focus on managing privileged access to data to reduce insider risks [11].

Ongoing high-profile breaches show further work remains to balance data use and privacy across regulatory environments. Organizations must navigate this landscape cautiously to earn customer trust.

V. INTERNET OF THINGS (IOT) AND CRITICAL INFRASTRUCTURE SECURITY

The massive expansion of connected Internet of Things (IoT) devices has introduced new attack surfaces for exploit. Many IoT devices lack basic security provisions, making them soft targets for botnets, DDoS attacks, and infrastructure infiltration. The 2016 Mirai botnet compromise of over 500,000 poorly secured IoT devices underscored these risks [12].

Industrial IoT and Operational Technology (OT) add layers of exposure in critical infrastructure sectors like energy, water, and manufacturing. Potential impacts range from service and production disruptions to threats to human safety. The 2021 Oldsmar water treatment facility hack—where attackers remotely increased lye levels—exemplified IoT risks to critical operations [13].

While sectors adapt standards like the NIST Cybersecurity Framework, IoT and ICS security lags, with progress largely reactive after major incidents. Securing these exponentially expanding attack surfaces remains an urgent concern for infrastructure resilience.

VI. CLOUD SECURITY

Migrating data, apps, and services to public cloud platforms provides efficiencies but also shifts security risks. The shared responsibility model gives cloud providers physical and environmental controls, while customers maintain data security, identity management, and application level controls. Insufficient cloud security posture leads to preventable data exposures.

Key cloud vulnerabilities include insecure interfaces, data breaches due to misconfigurations, unauthorized access to sensitive data, and vulnerabilities in serverless or container architectures [14]. Attackers exploit these gaps using stolen credentials, malware infections, and social engineering.

Recommended best practices for securing the cloud include enforcing least privilege access, data encryption, multi-factor authentication, continuous monitoring of configurations and activity, and regular audits. Cloud access security brokers like Microsoft Cloud App Security also provide visibility and control across cloud environments [15].

With appropriate security provisions, organizations can harness the cloud while avoiding the pitfalls of rapid adoption without adequate safeguards.

VII. HUMAN FACTORS IN CYBERSECURITY

Humans represent prime targets and unintentional gateways for determined attackers. Social engineering tactics manipulate trust and exploit human vulnerabilities to bypass technological controls. Well-crafted phishing emails often enable initial system infiltration.

Insider threats also loom as a top security concern, whether through intentional malfeasance or accidental human error. The 2022 IBM report found human error contributed to 62% of breaches, highlighting the need for robust cybersecurity awareness and training [16].

Education on spotting suspicious emails, practicing good password hygiene, identifying social engineering techniques, and reporting anomalies are vital for a strong human defence layer. Simulated phishing and ransomware training reinforce secure behaviours. With vigilance and tech-enabled assistance, people can become assets in thwarting cyber schemes.

VIII. REGULATORY AND POLICY FRAMEWORKS

Governments play a key role in defining cybersecurity standards and regulations that shape regional and industry security postures. In the U.S., pivotal policies include the Federal Information Security Management Act (FISMA), HIPAA security rules, and state breach disclosure laws. Government directives like Presidential Executive Order 14028 on "Improving the Nation's Cybersecurity" aim to strengthen federal and private sector security.

The EU Cybersecurity Act of 2019 expanded the cyber regulatory landscape with security requirements for critical infrastructure operators and vendors. The act also strengthened the European Union Agency for Cybersecurity (Enisa) to improve preparedness and resiliency [17].

However, regulations must balance security, privacy, transparency, and supporting innovation. Fragmented policy landscapes pose compliance challenges, while lagging laws fail to address modern technical realities. Close public-private sector collaboration can produce adaptable cybersecurity policies to meet evolving threats.

IX. CONCLUSION AND FUTURE DIRECTIONS

This analysis highlights the multifaceted terrain of cyber risks facing society today—from supply chain compromises to nation-state threats. As digitization accelerates across industries, the attack surface and stakes continue rising. Cyber resilient organizations prioritize network visibility, rapid response, and defence-in-depth leveraging people, processes and technology.

Looking ahead, imminent challenges include securing exploding IoT ecosystems, navigating the disappearance of the corporate perimeter, thwarting adversarial AI, and pushing international cyber norms and deterrence policies [18]. Taming rampant cybercrime requires global coordination among governments, enterprises, and technical communities. Cybersecurity necessitates continuous learning and persistence as we shape a trusted digital landscape for the future. With collaboration and vigilance, a more robust digital society lies within reach.

REFERENCES

[1] Gerwatowski, J. (2021). Cyberterrorism as an effect of the evolution of new information technologies in the first and second decades of the 21st century. World Complexity Science Academy Journal, 1(1), 297-323.

[2] Smith, B. (2016, October 21). Dyn statement on 10/21/2016 DDoS attack. Oracle Dyn. Retrieved from https://dyn.com/blog/dyn-statement-on-10212016-ddos-attack/

[3] Ehrenfeld, J. M. (2017, December 20). WannaCry, NotPetya, and the new age of cyber warfare. Harvard Business Review. Retrieved from https://hbr.org/2017/12/wannacry-notpetya-and-the-new-age-of-cyber-warfare

[4] Leventhal, R. (2019, April 15). A look back at the 2018 Marriott data breach. TechRepublic. Retrieved from https://www.techrepublic.com/article/a-look-back-at-the-2018-marriott-data-breach/

[5] Bing, C., Kerr, D., & Stubbs, J. (2020, December 14). Hackers used SolarWinds' dominance against it in sprawling spy campaign. Reuters. Retrieved [insert date], from https://www.reuters.com/article/global-cyber-solarwinds/hackers-used-solarwinds-dominance-against-it-in-sprawling-spy-campaign-idUSKBN28O2N8

[6] Lewis, J. A. (2018). Economic impact of cybercrime: No slowing down. Center for Strategic and International Studies (CSIS).

[7] Darktrace. (2019). Cyber AI: Response and resilience in the face of growing threats. Retrieved from https://go.darktrace.com/rs/478-REE-999/images/Report-Cyber-AI-Response-and-Resilience.pdf

[8] Recorded Future. (2019). Security Intelligence Handbook Chapter 6: Threat Intelligence. Recorded Future. https://go.recordedfuture.com/hubfs/ebooks/security-intelligence-handbook/Security_Intelligence_Handbook_Chapter_6.pdf

[9] Ayyagari, A. (2018, April 2). Demystifying blockchain: use cases in cybersecurity. Recorded Future. Retrieved March 6, 2022, from https://www.recordedfuture.com/blockchain-cybersecurity-use-cases/

[10] Tankard, C. (2016, June 9). GDPR compliance challenges and solutions. Digital Guardian. Retrieved from https://digitalguardian.com/blog/gdpr-compliance-challenges-and-solutions

[11] Ponemon Institute. (2020). Managing cyber risks in an IoT world. Retrieved from https://www.ponemon.org/local/upload/file/Cyber%20Risk%20in%20IoT%20World%20v3.pdf

[12] Antonakakis, M., April, T., Bailey, M., Bernhard, M., Bursztein, E., Cochran, J., ... & McCoy, D. (2017). Understanding the mirai botnet. In 26th {USENIX} security symposium ({USENIX} Security 17) (pp. 1093-1110).

[13] Paul, K. (2021, February 9). Oldsmar water treatment plant: How vulnerable is US infrastructure to cyber attack? The Guardian. Retrieved from https://www.theguardian.com/us-news/2021/feb/09/oldsmar-water-treatment-cyber-attack-us-infrastructure

[14] SANS Institute. (2019). Top cloud security threats and risks. Retrieved March 7, 2022, from https://www.sans.org/white-papers/40827/

[15] Microsoft. (2020). Microsoft Cloud App Security. Retrieved March 7, 2022, from https://docs.microsoft.com/en-us/cloud-app-security/

[16] IBM. (2022). Cost of a data breach report 2022. IBM Security. Retrieved from https://www.ibm.com/downloads/cas/OJDVQGRY

[17] Chivot, E., & Castro, D. (2019, April 17). What the EU gets right on cybersecurity. Center for Data Innovation. Retrieved from https://datainnovation.org/2019/04/what-the-eu-gets-right-on-cybersecurity/

[18] World Economic Forum. (2020). The Global Risks Report 2020. World Economic Forum. http://www3.weforum.org/docs/WEF_Global_Risk_Report_2020.pdf

Research on E-commerce Customer Value Segmentation Considering the Potential Value

Jinyi Huang and Ou Liu

***Abstract*—With the rapid growth of e-commerce platforms, enterprises face intense competition and must focus on existing customers to explore their value. Customer segmentation aids in formulating marketing strategies by identifying and categorizing customers with similar traits. The widely used RFM model characterizes customer value but solely considers current customer value, neglecting potential value. Hence, this study incorporates behavioral indicators of e-commerce users for segmentation. The unsupervised k-means++ algorithm is employed for clustering, with excellent results indicated by the silhouette coefficient. This analysis identifies a customer segment with high potential value, prompting enterprises to strengthen marketing strategies and convert these customers into loyal customers.**

***Index Terms*— e-commerce, customer value segmentation, RFM model, k-means++ algorithm.**

I. INTRODUCTION

Since the 21st century, the electronic commerce industry has experienced rapid development, with online shopping becoming one of the common channels for consumers. In 2022, global online retail sales are projected to reach $57.17 trillion, exhibiting a year-on-year growth of 9.71% [1]. Simultaneously, the continuous expansion of the online shopping market has led to unprecedented competition among e-commerce enterprises. However, the growth rate of online shopping users is gradually slowing down [2]. Consequently, e-commerce companies should focus on managing their existing customer base. The challenge they face is how to conduct precision marketing targeting their desired customers while minimizing marketing costs. Customer segmentation is a powerful method within customer relationship management, involving the identification and categorization of customers based on their attribute characteristics. Value segmentation is an effective customer segmentation method. To identify different levels of customer value, the RFM analysis model was proposed by Hughes, where RFM stands for Recency, Frequency, and Monetary [3]. However, when applying the RFM model for value-based customer segmentation in e-commerce, only the current value of customers based on their historical consumption behavior is considered. Potential value refers to the prospective profits that a customer may bring to a business in the future. Furthermore, potential value can reflect aspects of customer interest, needs, loyalty, and more, which can be inferred from various types of customer behavioral information. Currently, there is a lack of systematic research on customer value segmentation in the context of e-commerce. Therefore, this study aims to investigate how to segment e-commerce customers based on their potential value.

All authors are with the International Business School Suzhou, Xi'an Jiaotong-Liverpool University, Suzhou, China. email: Ou.Liu@xjtlu.edu.cn; katherine.jyhuang@gmail.com.

II. LITERATURE REVIEW

A. Customer Value Segmentation

Customer segmentation is an effective approach to categorize customers based on their distinct needs. Much of the research on e-commerce customer segmentation has emphasized customer value segmentation, with many studies enhancing the traditional RFM model for improved segmentation. One key strategy for enhancement involves the introduction of new metrics. In previous research, the periodicity of customers' visits was introduced into their model, and higher loyalty to the platform correlates with a greater likelihood of engaging in consumption behavior [4]. Some studies propose that customer value should encompass the entirety of a customer's lifetime profits, including both current and potential value, reflecting past and future profit contributions. Customer behavioral data such as the number of items added to the shopping cart and the number of items favorited serves as a novel source for customer segmentation [5]. Furthermore, following the definition of potential value, it can also represent a customer's potential to broaden their category of purchases and signify their demand structure [6]. Another strategy involves assigning weights to each metric, assessing the importance of these metrics. While Hughes also recognizes the significance of these three variables, numerous studies indicate that their importance varies depending on industry-specific characteristics.

Based on the above literature review of customer value segmentation, there are still some gaps in the existing literature. In the historical research, although the behavior data and the behavior of the product category have been introduced in the customer value segmentation research to determine customers' potential value, they have not been taken into consideration at the same time. Hence, it is impossible to determine their contributions to value segmentation respectively, and the potential value of customers based on user behavior could not be systematically discussed. Therefore, this paper makes innovations based on the traditional RFM customer value

segmentation and the detailed descriptions of the indicators are shown in Table 1.

TABLE I. E-COMMERCE CUSTOMER VALUE SEGMENTATION INDICATORS

Objective	Primary Indicators	Secondary Indicators	Third Indicators (Features)	Meaning
e-commerce customer value	Current Value	Purchasing power	Average Monetary (M)	Purchasing power
	Potential Value	Behavioral data	Purchase Frequency (F)	Activity
			View Frequency (V)	Activity
			Cart Frequency (C)	Activity
		Behavioral Data on Commodity Categories	Number of Product Categories Purchased	Dependence on the platform
			Number of Product Categories Viewed	Range of Interest
			Number of Product Categories Added to Cart	Range of Demand
			Number of Product Categories Interacted (including view, purchase and cart)	Lateral Activity
		Purchase Potential	View to Buy Conversion Rate	Purchase Potential
			Cart to Buy Conversion Rate	Purchase Potential (with clearer purchase intention)

B. Data Mining and Clustering

The most commonly used customer segmentation technique in data mining is cluster analysis. Many studies utilize the k-means algorithm, with some adopting enhanced versions. In a specific study, both k-means++ and EM algorithms were employed for customer segmentation, and the results indicated that, compared to other clustering algorithms, k-means++ achieved a higher level of accuracy [7]. Therefore, in this paper, the k-means++ algorithm is employed for customer segmentation.

III. METHODOLOGY

The methodology mainly includes four steps, which are 1. EDA and data preprocessing; 2. Feature engineering; 3. Index weight analysis; 4. Model fitting. To better visualize the process of methods, the research framework is shown in Figure 1.

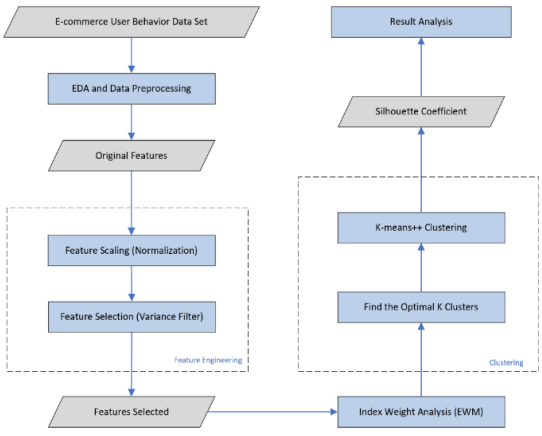


Fig. 1. Research Framework

A. Data

The dataset utilized in this study is a publicly available dataset sourced from Kaggle. It comprises customer behavior data collected over a three-month period, spanning from October 2019 to December 2019. This dataset encompasses a total of 30,933,357 records of customer interactions and involves 347,118 users from a sizeable multi-category e-commerce platform.

B. Index Weight Analysis

To ascertain the significance of each indicator within the enhanced model, this study employs Entropy Weight Method (EWM) to adjust the weights of these indicators, aiming to achieve more precise segmentation outcomes. The process of EWM encompasses multiple sequential steps.

Firstly, calculate the proportion P_{ij} of X_{ij}.

$$P_{ij} = \frac{Y_{ij}}{\sum_{i=1}^{n} Y_{ij}} \tag{1}$$

Secondly, compute the entropy value e_j of the j[th] index.

$$e_j = -\frac{1}{\ln(n)} \sum_{i=1}^{n} P_{ij} * \ln P_{ij} \tag{2}$$

The final step is to obtain the weight of each index.

$$W_j = \frac{1-e_j}{n-\sum e_j} \tag{3}$$

C. Clustering Using k-means++

The detailed steps of k-means++ are as follows:

1) Initialize the cluster centers.
   a) Randomly select a sample point as the initialized cluster center in the data set, which is described as X.
   b) Calculate the distance from each sample to the current cluster center and select the sample with the greatest distance as the next cluster center $C_k = x' \epsilon X$ based on probability $\frac{D(x')^2}{\sum_{x \epsilon X} D(x)^2}$, where D(x) represents the shortest distance between a data point x to the selected center.
   c) Repeat process b to determine all the cluster centers.
2) Calculate the distances of each data point from the centroids chosen in Step 1 using the Euclidian distance.
3) Compare the distance between each data point X_i and cluster centers C_k and assign each X_i to its closest cluster center C_k.
4) Calculate the mean of all X_i that have been assigned to the closest cluster center and update each cluster center as $\frac{1}{C_k} \sum_{x_i \epsilon C_k} X_i$.
5) Repeat steps 2, 3, and 4 until reaching the maximum number of iterations T, or when the difference between the two iterations is less than a certain threshold, the iteration terminates and receives the final clustering result.

IV. DATA PROCESSING AND MODEL FITTING

A. EDA and Data Preprocessing

Illustrated in Figure 2, page views constitute the most significant portion, representing roughly 90% of the total user interactions. In Figure 3, the average prices of products bought by users are predominantly situated below 500 yuan, with a focal concentration between 100 yuan and 200 yuan.

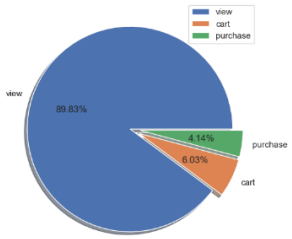


Fig. 2. User Behavior Distribution Diagram

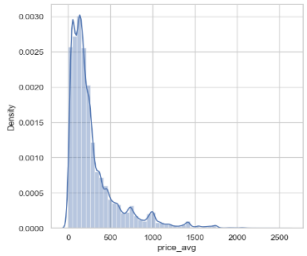


Fig. 3. Average Price Distribution of Purchases

Following EDA and data cleaning procedures, the data frame, initially organized with the event as the index, is transformed into a user behavior statistics table with user IDs serving as the index. Subsequently, the next stage involves the identification of noisy data, utilizing the Z-score function within SciPy. This process results in the inclusion of 335,090 users participating in the clustering analysis.

B. Feature Engineering

The initial step involves data normalization, followed by feature selection. Given the prior feature scaling, an estimator with a minimum threshold of 0.005 is employed for variance-based filtering. Features with variances falling below this threshold are eliminated. Ultimately, six features are chosen, and the selected features are presented in Table 2.

TABLE II. SELECTED FEATURES

Frequency (purchase)	Monetary (purchase_avg)	purchase_category	interactions_cateogry	view_buy	cart_buy

C. Indicator Weight Analysis

Each user's six indicators were multiplied by their respective weights to derive the weighted dataset.

D. Clustering Based on k-means++

The Davies-Bouldin Index (DBI) method is employed to determine the optimal number of K clusters in this research. As depicted in Figure 4, when K equals 4, the DBI is at its minimum, indicating that the clustering results are most favorable at this value.

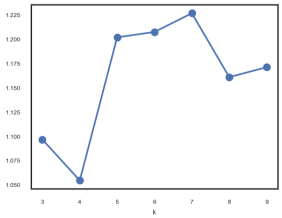


Fig. 4. Davies-Bouldin Index

Furthermore, the users are divided into four clusters, and the cluster centers are presented in Table 3.

TABLE III. CLUSTER CENTER

Feature / Cluster	Frequency	Monetary	purchase_category	interactions_cateogry	view_buy	cart_buy
1	0.006221	0.012973	0.004172	0.007882	0.011826	0.021767
2	0.011465	0.059013	0.005815	0.005328	0.015766	0.020108
3	0.157869	0.027113	0.074692	0.022298	0.015823	0.016960
4	0.049272	0.019826	0.040215	0.017550	0.011386	0.016579

Finally, the Silhouette Coefficient is employed to assess the clustering quality, with values ranging from -1 to 1. A higher Silhouette Coefficient indicates more cohesive clusters and greater separation between clusters, with values exceeding 0.50 indicating a favorable outcome [8]. In this study, the Silhouette Coefficient for the clusters is 0.53, indicating a good result.

V. RESULT AND DISCUSSION

A. Customer Value Segmentation Result

The result analysis employs a method that involves comparing the cluster centers with the dataset's average values. Subsequently, the current value, potential value, and Customer Lifetime Value (CLV) of each customer cluster are computed. Initially, the average values of the weighted results for each feature are calculated. Then, a comparison is made between the cluster centers and the average values. An upward-pointing arrow indicates that the cluster center surpasses the average value, while a downward-pointing arrow signifies that it falls below the average value. The outcomes are presented in Table 5.

TABLE IV. COMPARISON BETWEEN CLUSTER CENTERS AND AVERAGE VALUES

Feature / Cluster	Frequency	Monetary	purchase_category	interactions_cateogry	view_buy	cart_buy
1	↓	↓	↓	↓	↓	↑
2	⇓	⇑	↓	⇓	⇑	⇓
3	↑	↑	↑	↑	↑	↓
4	⇑	⇓	↑	⇑	⇓	⇓

Additionally, a comparison is made between the current and potential values of each cluster. Among the six selected features, ‘Frequency’ and ‘Monetary’ represent the current value of customers, while others belong to the customers’ potential

value. It should be emphasized that the lower the conversion rate, the higher the potential value. Thus, the conversion rates are negative when calculating potential value. The Customer Lifetime Value (CLV) is the sum of current and potential values, calculated using the following formulas.

$$Current\ Value = Frequency * W_{Frequency} + Monetary * W_{Monetary} \quad (4)$$

$$Potential\ Value = Purchase\_category * W\_(Purchase\_category) + Interactions\_category * W\_(Interactions\_category) - View\_buy * W\_(View\_buy) - Cart\_buy * W\_(cart\_buy) \quad (5)$$

$$Customer\ Lifetime\ Value\ (CLV) = Current\ Value + Potential\ Value \quad (6)$$

Where $W_{Frequency}$ represents the weight of the indicator Frequency calculated by the EWM. And so on for the rest of the features. The results and the profile tags of customer clusters are shown in Table 6.

TABLE V. CURRENT VALUE, POTENTIAL VALUE AND CLV

Value / Cluster	Current Value	Potential Value	CLV	CLV Ranking	Customer Profile
1	0.019193	0.021539	0.002346	4	Ordinary Customers
2	0.070478	0.024731	0.045747	3	Customers with Strong Purchasing Power
3	0.184982	0.064207	0.249189	1	Active Customers
4	0.069098	0.029801	0.098899	2	High Potential Value Customers

B. Discussion of Clustering Results

Finally, detailed analysis and marketing strategies for each customer cluster are carried out.

1) Cluster 1: Ordinary Customers. This group constitutes 63.9% of the total customer base, which is the largest customer group in this study. Characteristics of this consumer group include the lowest current value and potential value. Therefore, such customers with lowest CLV are classified as ordinary users, and e-commerce companies can adopt an observant approach towards them.
2) Cluster 2: Customers with Strong Purchasing Power. This group comprises 13.3% of the total, characterized by strong purchasing power but low buying frequency, ranking second in terms of current value. However, they show a relatively limited range of interests according to the potential value. Considering both current and potential value, their CLV ranks third, higher than cluster 1. Therefore, such customers can be classified as customers with strong purchasing power and should be maintained by the e-commerce company.
3) Cluster 3: Active Customers. This customer category represents the smallest proportion of the total customer base, accounting for only 4.4%. They are characterized by high purchase frequency, followed by strong purchasing power, and they hold the highest current value. Notably, these customers exhibit both high current and potential values. In summary, this customer segment boasts the highest CLV, making them the most profitable for the business among all groups.
4) Cluster 4: High Potential Value Customers. The proportion of such customers is 18.5%, which is relatively large. They rank third in current value but exhibit relatively high potential value. Specifically, they engage with a wide variety of product categories, albeit with a lower conversion rate, indicating a broad willingness to make purchases on the platform and maintain a certain level of attention to the e-commerce platform. Based on the current and potential value, this type of user ranks second in CLV and is a key development target for the enterprise. Therefore, this type of customer needs to be vigorously developed among enterprise customers. The company can allocate appropriate marketing investments to stimulate and increase customer frequency, thus enhancing customer loyalty and substantially boosting the company's profits.

In summary, active customers (Cluster 3) and high purchasing power customers (Cluster 2) hold the highest current value. In terms of potential value, active customers still have the highest potential value, followed by high potential value customers (cluster 4). Ordinary users (cluster 1) display the lowest existing and potential value. From the perspective of CLV, active customers undeniably hold the top position. The second place in CLV ranking is the high potential value customers, which is the most notable result of this paper. Furthermore, the e-commerce platform can allocate additional resources to these customers, with the aim of converting their potential value into actual value and thereby increasing profits for the enterprise.

VI. CONCLUSIONS

This paper studies the value segmentation indicators and value segmentation method based on the e-commerce customer behavior data, ultimately achieving customer segmentation and assisting enterprises in formulating their marketing strategies. The main contributions are improving customer value indicators, improving the customer segmentation method and establish a more reasonable value segmentation index system and receiving remarkable results of customer value segmentation through k-means++ clustering. Among four clusters resulted from the customer value segmentation, the biggest contribution of the improved model is to identify customers with high potential value although their current value ranks third. The e-commerce platform can develop marketing strategies to stimulate these customers' potential value. In addition, this paper also analyzes the characteristics of each cluster and provides recommendations to help e-commerce platforms achieve successful customer relationship management.

REFERENCES

[1] Statista. Retail e-commerce sales worldwide from 2014 to 2026 (in billion U.S. dollars). URL: https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/, last check on 2023-11-26.

[2] Marketer. Global Ecommerce Forecast 2022. URL: https://www.insiderintelligence.com/content/global-ecommerce-forecast-2022, last check on 2023-11-26.

[3] A. M. Hughes. Boosting response with RFM. *Marketing Tools*, 3(3), pp. 4.

[4] S. Peker, A. Kocyigit, and P. E. Eren. LRFMP model for customer segmentation in the grocery retail industry: a case study. *Marketing Intelligence and Planning*, 35(4), pp. 544-559.

[5] R. Zhao, and C. Li. Research on E-commerce Customer Segmentation Based on RFAC Model. *2021 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS)*, pp. 439-444.

[6] P. C. Verhoef, B. Donkers. Predicting customer potential value an application in the insurance industry. *Decision Support Systems*, 32(2), pp. 189-199.

[7] F. Yoseph, N. Malim, M. Heikkilä, and A. Brezulianu. The impact of big data market segmentation using data mining and clustering techniques. *Journal of Intelligent & Fuzzy Systems*, 38(5), pp. 6159-6173.

[8] E. Mooi, and M. Sarstedt. A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistic. *Springer Berlin Heidelberg*.

Integrating Virtual Reality into Classroom-as-Organization Learning Design: Experimental Case Study

Dr Jean-Yves Le Corre and Qinyi Huang

***Abstract*— Several experts acknowledge that constructivist learning strategies and instructional design goals should be supported by clear instructional design frameworks and guidelines. However, educators may find it difficult to find practical methodologies or frameworks when it comes to designing and implementing immersive learning experiences. This case study aims to create and integrate Virtual Reality digital content into an existing course in a Learning Management System. The objective is to identify Virtual Reality digital contents of different forms to serve different purposes, frequencies, and modes of integration for a course designed based on the Classroom-as-Organization learning model.**

Keywords*— *Constructivist Learning, Virtual Learning Environment, Immersive Learning, Virtual Reality

I. INTRODUCTION

From the early stages of the development of constructivist learning strategies, using advanced digital technologies, experts have claimed that this would bring new possibilities to the design and implementation of immersive learning solutions. However, educators have to encounter several challenges in integrating VR digital content into learning design, when creating and integrating VR digital content in a course designed and implemented in a Learning Management System (LMS).

Literature suggests that one reason for the limited implementation of immersive virtual environment applications may be a lack of recommendations for their effective design [3]. While scholars have acknowledged a dearth of practical guidelines for learning design in immersive environments, there exist only a few attempts to bridge this gap. Workspaces in 3-D virtual spaces could represent other attractive opportunities to create immersive learning experiences. Guerrero [1] found that the design of a Collaborative Virtual Environment (CVE) for people who know each other and interact in a real space is different from the traditional CVE design and that designing virtual environments to support collaborative work may present several advantages for collaborative learning.

The approach adopted in this study is to create and integrate VR digital content into an existing course based on the Classroom-as-Organization (CaO) instructional design model. VR digital content may have different forms and serve different purposes, frequencies, and modes of integration.

All authors are with Xi'an Jiaotong-Liverpool University. Emails: Jeanyves.Lecorre@xjtlu.edu.cn, Qinyi.Huang22@student.xjtlu.edu.cn

II. INSTRUCTIONAL DESIGN

Several experts acknowledge that constructivist learning strategies and instructional design goals should be supported by clear instructional design frameworks and guidelines.

A. Classroom-as-Organisation

The Classroom-as-Organisation (CaO) instructional model [2] was examined to identify learning design components across three strata: the learning environment, the learning process, and performance tasks. The objective of this model is to recreate organizational contexts where learners are immersed in an organization and typical situations at the workplace, involving interactions with various fictitious roles for collaborative learning and knowledge construction.

B. Learning Prototype

A prototype named ‘Virtual Social Entrepreneurship Hub’ was developed in an LMS (Moodle). The Virtual Social Entrepreneurship Hub is a virtual space where student consultants work collaboratively with small business owners to help them start, develop, and grow their business ideas. The Virtual Social Entrepreneurship Hub connects student consultants with business owners, entrepreneurship mentors, business experts, and teachers. The role of the teacher is to assist students-consultants in the preparation of a pre-evaluation of the business idea called the 'Business Road Test' Student-consultants, business owners, and mentors can enter the Virtual Hub One-Stop Counter to monitor their progress in the completion of the different steps for preparation of the 'Business Road Test'. Virtual Hub One-Stop Counter provides a roadmap to student consultants for the completion of the 'Business Road Test' containing a list of synchronous and asynchronous activities. Synchronous learning refers to instructors and students gathering at the same time and (virtual or physical) place and interacting in “real-time”. Asynchronous learning refers to students accessing materials at their own pace and interacting with each other over longer periods. At the

synchronous end of the spectrum lie activities such as live-streaming lectures and participating in video-conference discussions. At the asynchronous end lie activities such as watching pre-recorded lectures, reading assigned materials, and participating in discussion boards. Mentors give a lot of target opinions and suggestions about how to improve current conditions for business owners. Finally, students can get into business network forums to connect counselors, business owners, and business executives in small and larger firms through business networks. GBA research services can serve as an aid to provide easy access to student consultants to collect market and industry data and information from public sources.

C. Research Questions

The study aims to address several critical issues that may arise when integrating VR digital content for Classroom-as Organizations (CaO) learning design including:

- What are the main functions of VR technologies able to support the learning objectives of a CaO-based course?
- How would VR digital content fit into learning activities for Classroom as Organizations?
- To what extent would VR technology be able to enhance the learning experience?

The different types of media content and VR content may depend on the type of VR technology platform used to create the immersive learning experience:

- What kind of synchronous learning activities could be fully or partially supported by the VR technology platform?
- How would the graphic interface in the VR platform (either through the web or VR headsets) support those learning activities?
- Which types of media and/or VR digital contents e.g., full-size 3D models, Interactive 360° images, and 360° videos, 2D pictures, images, audio files…

III. IMMERSIVE LEARNING PROTOTYPE

A. Real-life scenarios

Students as consultants can work together with small business owners to help owners start, develop, and grow their business ideas while several scenarios and storyboards are created to support this process, along with virtual reality content to support the immersive learning experience.

Fig. 1. Operation Interface of Virtual Room in Nibiru

The teacher plays the role as coordinator and coordinates the preparation of the “Business Road Test” for the project with the business owner that has been assigned to student consultants. Then students can meet with business owners to grasp business ideas, connectedness across the value chain, and others about business owners. Under the guidance of the teacher, students can meet with mentors who are experienced in startup entrepreneurship.

It was decided to upgrade the course design and create some new characters which could be included in the videos. The new roles included those of facilitator, coordinator, and supervisor of every activity. They give a brief introduction to students, leading students into a virtual discussion room to have more in-depth communication with teachers, mentors, and figures in the business networks forum. They could come back to the virtual hub one-stop counter after they finish part of the discussion and enter the next scenery. The speeches, characters, and shooting locations should all be taken into consideration in the learning design.

B. Design Steps

Researchers used a special camera to shoot 3D videos to simulate the real scenery of the Virtual Social Entrepreneurship Hub. Before shooting, they prepared for storyboarding brainstorming and making templates of speeches. Seven scenes were shot, they were the entrance of the incubator; reception desk; transit to the virtual hub one-stop counter, teacher’s room, mentor’s room, business networks forum, and GBA research services respectively.

Nibiru creator can create a large VR immersive teaching classroom, which can break into the physical space limitations of traditional teaching, concretely present knowledge points, solve teaching key and difficult points, and make it easy to understand. It can also realize scene teaching and strengthen deep memory through immersive experience.

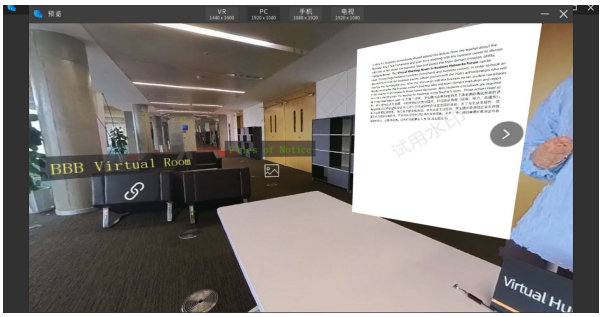


Fig. 2. Preview of Instructions and Learning Activities in Nibiru

The 3D pictures and videos can be imported into Nibiru work template. Some scene-switching buttons can be used to change from place to place. There are real avatars who can make introductions to different meeting rooms and lead students in first perspective into different rooms to have in-depth discussions. Some attachment files and learning materials can be shown under trigger clicking. Also, students from the first perspective can enter virtual rooms to talk about many kinds of roles in the entrepreneurship hub. The links of scenarios can be inserted into the LMS platform windows to showcase immersive, real-time interactive, and more comprehensive experiences. When students in the first perspective click the rooms in the LMS, they can enter virtual scenes to have an in-depth experience.

C. Virtual Reality Software

Nibiru Creator can render the scene completely on a flat plane, and it can play 3D videos with high interactivity and integrate Nibiru content in the course design. For the learning environment, Nibiru can create immersive learning spaces, virtual (work) spaces, and navigation between virtual (work) spaces, using 3D videos with real avatars that have already been shot. For the learning process, Nibiru can create hyperlinks to generate entrance into virtual rooms for discussion, and there are many-branched scenarios. Students can have synchronous and asynchronous activities in the Nibiru setting. When teachers and students come together at the same time and location (physical or virtual), they are said to be engaged in "real-time" interaction. Asynchronous learning is when students communicate with each other and access resources at their own pace over a longer period.

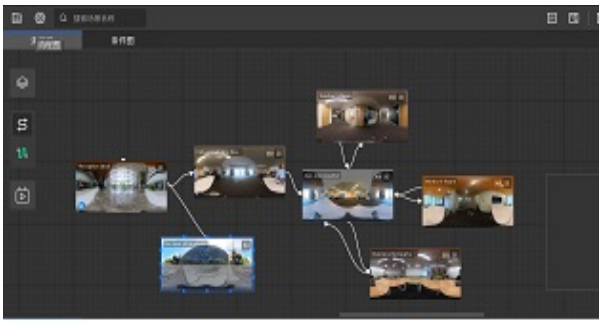


Fig 3. Overview of Project Construction in Nibiru

IV. CONCLUSION

The table below provides an overview of the different types of VR digital contents and mode of integration into the course in the LMS.

SVEH (*)	Type of VR digital contents		
	Learning Environment	Learning Process	Performance Tasks
Integration with Moodle (LMS)	It is a platform where different rooms are presented, and students can click them to enter and have a brief understanding.	Students can read instructions and learning materials from different sections in the Social Virtual Entrepreneurship Hub.	Students can receive comments and report idea summaries.
VR digital contents (Nibiru)	It has several 3D scenes where real settings can be simulated.	Students can know the next steps under the guidance of coordinators in transit scenarios.	Students can enter the virtual BBB room to have in-depth discussions with characters in SVEH.
Frequency of learning activity	It may take about 20 minutes for students to be familiar with single-room content.	It may take about 30 minutes for students to grab related knowledge of one part.	It may take at least 60 minutes for students to meet and interact with characters online.

*Social Virtual Entrepreneurship Hub.

ACKNOWLEDGMENT

We would like to thank the company Nibiru for granting access to the Nibiru Creator software during the period of completion of this study.

REFERENCES

[1] L. Guerrero, C. Collazos, J. Pino, S. Ochoa, and F. Aguilera, "Designing Virtual Environments to Support Collaborative Work in Real Spaces," J. Web Eng., vol. 2, pp. 282-294, 2004.

[2] D. R. Thomas, S. F. Chappell, and D. S. Bright, "Classroom as Organization," Edward Elgar Publishing, Cheltenham, UK, 2020. [Online]. Available: https://doi.org/10.4337/9781788979856. Accessed: Aug. 6, 2023.

[3] T. T. De Back, A. M. Tinga, and M. M. Louwerse, (2021): Learning in immersed collaborative virtual environments: design and implementation, Interactive Learning Environments, 2021 [Online]. Available:https://doi.org/10.1080/10494820.2021.2006238

Stability measurement on rowing movement

Sheng-Ru Hsiao, Chyuan-Huei Thomas Yang, Chien-Chang Chen and Cheng-Shian Lin

***Abstract*—Rowing is a sport that requires players to demonstrate high-density coordination and consistency in their movements. Stable and consistent rowing strokes by players lead to better performance. This study analyzes rowing movements recorded from rowing machine to measure the stability of single player. The proposed method first uses the OpenPose system to extract body joint poinst from rowing video. The joint point information was then fed into the Temporal Cycle Consistency (TCC) learning network, named as Pose Temporal Cycle-Consistency (PTCC), for training the model. The proposed PTCC network aligned the timelines of the videos as closely as possible and generated frame-by-frame retrieval results. At last, the average matched frames among each cycled rowing video represents the players rowing stability. Experimental results show that using some players' rowing videos can train the PTCC model and the similarity calculated from average matched frame rates is efficient to determine the stability of one player.**

***Index Terms*—OpenPose, Contrastive Learning, Temporal Cycle-Conslstency Learning, Frame by frame retrieval.**

I. INTRODUCTION

In the long-standing development of research in computer vision, the topic of human body pose and motion has been a frequently explored area, often applied in sports-related research. However, obtaining accurate joint coordinates of the human body has been a challenging task in past studies. Many methods have been developed to address this issue. One such approach involves attaching markers to different parts of the body to detect joint points and this method is susceptible to generating noise due to the friction caused by the markers on the human body during movement [1]. Therefore, another type of wearable device was specifically designed. This device can acquire relevant data through sensors placed all over the body [2]. Compared to other methods, this device offers advantages such as increased precision in obtaining human joint coordinates. However, since physical devices have a certain weight and size, they can easily influence the player's mobility and cause some interferences on their physical activities, potentially causing disruptions in their movements and leading to errors in experimental data. The second one uses the Microsoft Kinect to detect human joint points [3]. The Kinect Sensor consists of a depth sensor, a color camera, and a four-microphone array, providing capabilities for full-body 3D motion capture, facial recognition, and speech recognition. The third approach involves the use of the Nintendo Switch [4]. The Nintendo Switch combines the concepts of a home console and a portable gaming device. It incorporates motion control devices built into the two Joy-Con controllers, an infrared camera, and other peripheral accessories. These devices enable motion capture for both wearable and handheld applications. While Kinect and Switch were originally developed for home gaming consoles, they have also influenced the direction of research in the detection of human joint points.

With the passage of time and the advancements in technology, devices such as smartphones, GoPros, drones, digital cameras, and more not only make it effortless to capture high-quality video but also provide the convenience of portability. Nowadays, the integration of artificial intelligence applying on various sports videos enables coaches to analyze them more efficiently. Two well-known human pose estimation systems, named as the OpenPose system developed by Carnegie Mellon University (CMU) [5] and the MediaPipe system developed by Google Research [6], are presented recently. Both systems can detect the coordinates of human joint points in videos and evaluate human posture accurately. This feature benefits researchers in subsequent sports research applications and eliminates the need of participants to wear any additional equipment. After examining the properties among OpenPose and MediaPipe, this study adopts pose information detected by the OpenPose for further rowing analysis.

Recently, unsupervised learning techniques have demonstrated their superior ability to prepare training data without labeling. Therefore, the study of unsupervised learning techniques for training large amounts of rowing videos without labeling is worth investigating. This study adopts the model of contrastive learning, which analyzes videos frame-by-frame to learn motion, for learning human body movements. The proposed method applies pose information to the contrastive learning model to learn the rowing motion. One player's rowing video is then applied to the learned contrastive learning model to measure the similarity of this player

The remaining sections of the paper are organized as follows. Section 2 reviews important related background. Section 3 introduces our proposed Pose Temporal Cycle-Consistency Learning (PTCCL) model and similarity measurement methods. Section 4 shows the experimental results of the proposed method. Section 5 concludes the proposed method.

II. RELATED BACKGROUND REVIEWS

Section II reviews important works related to this study. Section II.A introduces and compares two human pose recognition systems OpenPose and MediaPipe. Significant features are listed to explain the reasons of choosing OpenPose

Sheng-Ru, Hsiao, Department of Computer Science and Information Engineering, Tamkang University, Taiwan, 610410572@gms.tiki.edu.tw
Chyuan-Huei Thomas Yang, School of Information Engineering, Shandong Vocational and Technical University of International Studies, Rizhao, Shandong, 276826 P. R. China, chyang_0922412256@qq.com
Chien-Chang Chen, Department of Computer Science and Information Engineering, Tamkang University, Taiwan, ccchen34@mail.tku.edu.tw
Cheng-Shian Lin, Department of Computer Science and Information Engineering, Tamkang University, Taiwan, 157446@mail.tku.edu.tw

in our study. Section II.B introduces the video-based contrastive learning model.

A. Human Pose Recognition System

In the field of action recognition, there are two main approaches: image-based recognition using deep learning and action detection based on human pose estimation. Applications such as human-computer interaction [8], human behavior understanding [9], or medical assistance [10] fall within the realm of deep learning-based action recognition. However, these methods often involve large models, leading to suboptimal execution efficiency. On the other hand, traditional human pose estimation methods lack accuracy and real-time performance and struggle to operate effectively in complex environmental backgrounds [11,12,13]. Nonetheless, there are approaches that do not require additional wearable devices for joint point recognition. For example, systems like Kinect enable human pose estimation through software algorithms directly in videos. Hence, we opt for a human pose estimation system to construct a player's motion analysis system. To achieve better experimental results, we examine two different human pose estimation systems, OpenPose [5] and MediaPipe [6], to determine which one is more suitable for our system.

The OpenPose system was introduced by Carnegie Mellon University (CMU) in 2017 [5]. It is an open-source system available for use by all relevant researchers. Developed using C++ with OpenCV and Caffe, OpenPose is a multi-threaded, real-time, multi-person human pose estimation system. It can detect the coordinates of various joints in the human body, as well as facial and hand landmarks. As an open-sourced software, OpenPose is designed to be installed on various operating systems, including Windows, MacOS, and Linux.

MediaPipe is a multimedia cross-platform machine learning framework introduced by Google Research in 2019. MediaPipe primarily uses the BlazePose algorithm [14] and the COCO human body keypoints [15] as the basis for detecting the body skeleton and joint points. These capabilities allow developers to conduct research on specific body parts more flexibly.

B. *Temporal Cycle-Consistency Learning*

In 2019, Dwibedi et al. [16] presented another self-supervised learning method, Temporal Cycle-Consistency Learning(TCC), and applied to video analysis. The TCC is an improved method of the TCN. The objective of TCC is to learn mutual mappings between feature spaces of two different videos while preserving their temporal structures. The TCC utilizes the Cycle-Consistency Loss (CCL) function to map the correlations between two feature spaces. Fig .1 shows the diagram of TCC operation.

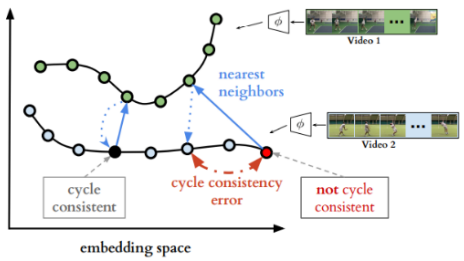


Fig. 1. TCC operation flow [16].

The TCC starts by taking two different videos with similar actions and feeding them into an image encoder to extract their features, forming as embedding spaces. The embedding sequences U and V are obtained by encoding video sequences S and T with the encoder network, respectively. For the selected frame i in U, each cycle of the TCC utilizes the soft nearest neighbor computation to identify the closer frame and then optimizes the frame as $\tilde{v}$ and cycling back to U again for finding the corresponding frame. Therefore, the searched highest peak frame is considered as the optimal matched frame to the original one. Fig. 2 depicts the TCC structure.

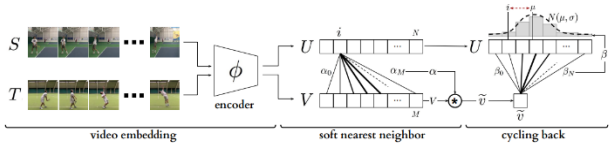


Fig. 2. The TCC structure [16].

III. PROPOSED METHOD

This section introduces our proposed stability detection method on a rowing video, that is composed of one minute rowing movement on the same player. The proposed scheme includes three procedures, including the rowing cycle segmentation (*RSC*), the adopted Pose Temporal Cycle-Consistency (*PTCC*) and the rowing stability measurement (*RSM*). This section introduces the steps of the proposed scheme.

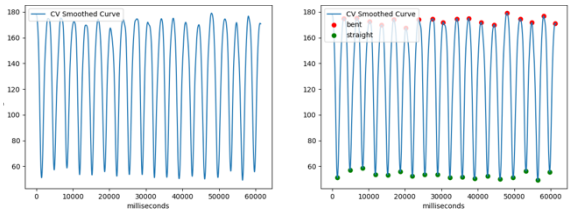
First, the proposed rowing cycle segmentation (*RSC*) procedure segments the one-minute rowing video to lots of cycled rowing video. For the joint points extracted from the OpenPose system, angle can be calculated from Eqs. (1) and (2).

$$radian = arctan\left(\left(\frac{y_3-y_2}{x_3-x_2}\right)-\left(\frac{y_1-y_2}{x_1-x_2}\right)\right) \quad (1)$$

$$angle = \left|\frac{radian\times 180.0}{\pi}\right| \quad (2)$$

where (x_1, y_1) , (x_2, y_2) , and (x_3, y_3) represent coordinates in x and y axes of three joint points. After the angles are calculated, a smooth function is applied to the angle change and Fig. 3(a) shows the smoothed curve of the selected elbow change. The local maximum of the smoothed curve determines start or end frame of each rowing cycle. Therefore, the movement change frame in rowing can be acquired like Fig. 3(b). Consequently, the one-minute rowing video can be

segmented to combination of one-cycled rowing video, Ci ($1 \leq i \leq n$) with totally n rowing cycles.



(a) (b)

Fig. 3. Smoothed elbow angle change curve.

Several players' rowing cycles Ci are then applied to the *PTCC* model, training the *TCC* model by using Pose data, in which each cycled video is segmented to combination of frames and then train the model by pose information of each frame. The *PTCC* model is trained by the following steps:

1. Segment all input videos to collection of one-cycled rowing videos by the proposed rowing cycle segmentation (RSC).
2. Detect human pose from frames of each rowing cycle.
3. Train the *TCC* model by all rowing cycles.

Fig. 4 depicts flowchart of training the adopted *PTCC* model. After the model is trained, the following rowing stability measurement (*RSM*), depicted in Fig. 5, are proposed to measure the stability of a rowing player.

1. Segment one player's rowing video to collection one-cycled rowing videos by the proposed rowing cycle segmentation (*RSC*).
2. Calculate the rowing similarity between one rowing cycle and other rowing cycles by the following steps:
   - I. Assume Ci denotes number of rowing cycle, in which $1 \leq i \leq |C|$ and $|C|$ is the cycle number of one player's video.
   - II. Calculate rowing similarity between Ci and other rowing cycles by Eq. (3).

$$Sim_i = \frac{1}{|C-1|}\sum_{j=1,j\neq i}^{|C|} Sim(C_i, C_j) \tag{3}$$

   - III. At last, the optimal stability measurement of rowing movement is acquired from Eq. (4).

$$\max\{Sim_i\}, 1 \leq i \leq |C| \tag{4}$$

The proposed *RSM* first calculates the similarities between one cycled rowing video and other rowing cycles. Then, the average of similarity between one cycled rowing video and other rowing cycles are calculated as shown in Eq. (3). At last, the maximum average similarity value is the stability of one player's rowing video.

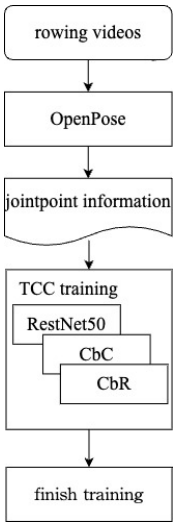


Fig. 4. Flowchart of the adopted Pose Temporal Cycle-Consistency (*PTCC*).

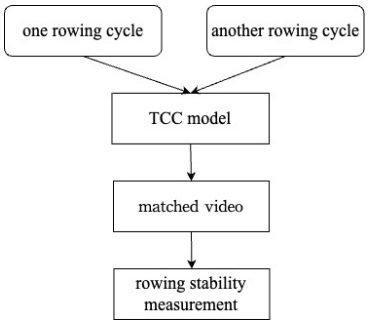


Fig. 5. Flowchart of the rowing stability measurement (*RSM*).

Since there are many rowing cycles in one player's rowing video, an improved method is presented to calculate the stability without exhausted comparing all pairs of rowing cycles.

Assume that the rowing video C includes n rowing cycles, as denoted by $C_i (1 \leq i \leq n)$, and each rowing cycle i includes $x_{i,j} (1 \leq j \leq |C_i|)$ frames. Eq. (5) defines the set of frames of the video C.

$$C = \{x_{i,j} | 1 \leq i \leq n, 1 \leq j \leq |C_i|\} \tag{5}$$

Then, the average frame number among these cycles is acquired from Eq. (6), in which $|C_i|$ represents the frame number of cycle C_i.

$$\bar{C} = \frac{\sum_{i=1}^{n} |C_i|}{n} \tag{6}$$

Since the rowing video C may include some rowing cycles with frame number far away from the $\bar{C}$, therefore, the new cycle set C_{new} is acquired from Eq. (7) by user defined range parameter τ.

$$C_{new} = \{C_i | \ ||C_i| - \bar{C}| \leq \tau\} \tag{7}$$

At last, the optimal frame number $\overline{C_{new}}$ from the new cycle set C_{new} is defined as Eq. (8).

$$\overline{C_{new}} = \frac{\sum_{C_i \in C_{new}} |C_i|}{|C_{new}|} \quad (8)$$

After the optimal frame number $\overline{C_{new}}$ is determined, those cycles with frame number equal to $\overline{C_{new}}$ are selected as the optimal cycles C_{select}. At last, for each cycle $C_s (C_s \in C_{select})$, its average matched frame rate m_s is calculated by the average matched rate between the cycle C_s and all other cycles in C.

IV. EXPERIMENTS

This section, including experimental dataset and experimental result, of the proposed rowing movement stability measurement. Section 4.1 introduces the experimental dataset and Section 4.2 shows the experimental results.

The experimental dataset is composed of 3 players' videos, in which each video is recorded by one minute around 1800 frames with resolution 1920×1080. Fig. 6 shows frame number of each cycle on player one. In this figure, frame number of each cycle of the player is ranged from 34 to 46. Although the figure shows that the curve is slowly climbing, a mathematical measurement between all cycles worth our study and the adopted *PTCC* and *RSM* methods are presented to measure the stability between cycles systematically.

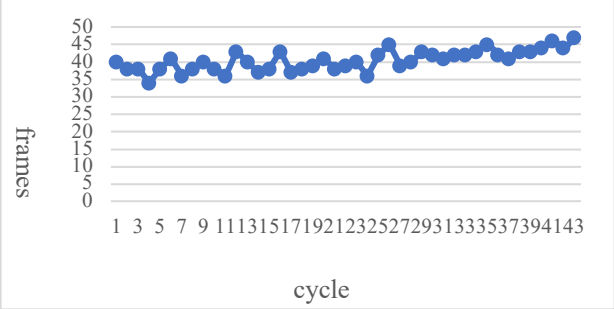


Fig. 6. Frame number of each rowing cycle on one player.

The *PTCC* model is trained by 1000 epochs, and the loss function of training and validation is depicted in Fig. 7. Fig. 7 also shows that there is no overfitting in the proposed *PTCC* model training. Moreover, the data trained after 1000 epochs performs more closer classified than before training.

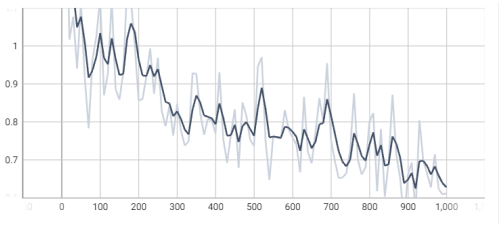


Fig. 7. Loss function of the adopted *PTCC* model.

This section shows experimental results of the proposed rowing stability measurement method. First, two different rowing videos can be applied to the trained *PTCC* model for acquiring the matched result. Fig. 8 shows the matched results of two similar directional rowing videos. In this example, two different players' videos are applied for generating the matched results. Figs. 8(a) and 8(b) are two sets of matched frames, in which Fig. 8(a1) and Fig. 8(b1) are matched, and so on.

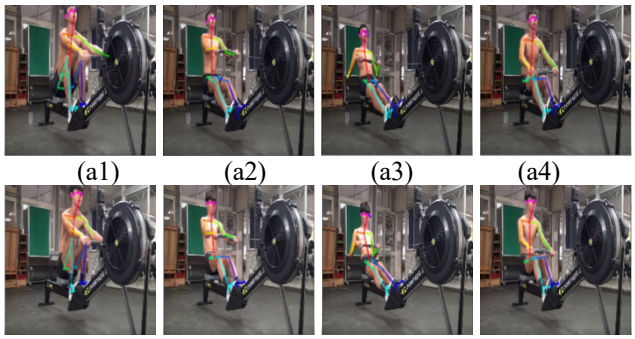


Fig. 8. Matched frames of two similar directional rowing video.

Table 1. Lists of average matched frame rate.

cycle number	match rate (%)
4th	93.3
8th	93.6
12th	**95.2**
27th	92.4

Table 1 shows that the 12th cycle outperforms the others. Moreover, the optimal match rate is 95.2%. Therefore, the rowing stability of the player is also defined as 95.2%. Table 2 lists average matched frame rates of another player. In this table, the maximum match rates of another player is 94.63%. These experimental results demonstrate the rowing stability of each player.

Table 2. Lists of match rates of another player.

Player	cycle number	match rate (%)
Player 2	7th	92.5
	11th	92.44
	12th	**94.63**
	16th	93.28
	17th	92.62
	18th	92.27
	22th	92.86

V. Conclusion

This study presents an efficient stability measurement method on rowing player. The proposed method involves three important procedures, the rowing cycle segmentation (*RSC*), the Pose Temporal Cycle-Consistency (*PTCC*), and the rowing stability measurement (*RSM*). The proposed RSC method segments the input one-minute video into collection of cycled videos. The adopted method, named Pose Temporal Cycle-

Consistency (*PTCC*) model, trains the unsupervised TCC model by using the pose information of the cycled video extracted from the OpenPose system. After the model is trained, the rowing stability measurement (*RSM*) acquires the stability by calculating the match rates among cycles of the rowing video. Experimental results show that proposed method can distinguish the rowing stability of one player among several players. The most stable rowing player can therefore be detected by using the unsupervised model.

REFERENCES

[1] Reinschmidt, C., Van Den Bogert, A. J., Nigg, B. M., Lundberg, A., Murphy, N., "Effect of Skin Movement on the Analysis of Skeletal Knee Joint Motion During Running," *Journal of biomechanics*, 30(7), pp. 729-732, 1997.

[2] Fujimori, Y., Ohmura, Y., Harada, T., Kuniyoshi, Y.,"Wearable Motion Capture Suit with Full-Body Tactile Sensors," In *2009 IEEE International Conference on Robotics and Automation*, pp. 3186-3193, May, 2009.

[3] Liu, L., Wu, X., Wu, L., Guo, T. (2012, October). "Static Human Gesture Grading Based on Kinect," In *2012 5th IEEE International Congress on Image and Signal Processing*, pp. 1390-1393, October, 2012.

[4] Dempsey, P., "The Teardown-Nintendo Switch Gaming System," *Engineering & Technology*, 12(4), pp. 82-83, 2017.

[5] Cao, Z., Simon, T., Wei, S. E., Sheikh, Y., "Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7291-7299, 2017.

[6] Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Grundmann, M., Mediapipe: A Framework for Building Perception Pipelines," *arXiv preprint arXiv:1906.08172*, 2019.

[7] Hadsell, R., Chopra, S., LeCun, Y., "Dimensionality Reduction by Learning an Invariant Mapping," In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 1735-1742, June, 2006.

[8] Ren, Z., Meng, J., Yuan, J., Zhang, Z, "Robust Hand Gesture Recognition with Kinect Sensor," In *Proceedings of the 19th ACM international conference on Multimedia*, pp. 759-760, November, 2011.

[9] Wei, S. E., Tang, N. C., Lin, Y. Y., Weng, M. F., Liao, H. Y. M., "Skeleton-Augmented Human Action Understanding by Learning with Progressively Refined Data," In *Proceedings of the 1st ACM International Workshop on Human Centered Event Understanding from Multimedia*, pp. 7-10, November, 2014.

[10] Huang, J. D., "Kinerehab: a Kinect-Based System for Physical Rehabilitation: a Pilot Study for Young Adults with Motor Disabilities," In *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*, pp. 319-320, October, 2011.

[11] De Smedt, Q., Wannous, H., Vandeborre, J. P., Guerry, J., Saux, B. L., Filliat, D., "3D Hand Gesture Recognition Using a Depth and Skeletal Dataset: Shrec'17 track," In *Proceedings of the Workshop on 3D Object Retrieval*, pp. 33-38, April, 2017.

[12] Devineau, G., Xi, W., Moutarde, F., Yang, J., "Convolutional Neural Networks for Multivariate Time Series Classification using both Inter- and Intra-Channel Parallel Convolutions," In *Reconnaissance des Formes, Image, Apprentissage et Perception* (RFIAP'2018), June, 2018.

[13] Hou, J., Wang, G., Chen, X., Xue, J. H., Zhu, R., Yang, H., "Spatial-Temporal Attention Res-TCN for Skeleton-based Dynamic Hand Gesture Recognition," In *Proceedings of the European conference on computer vision (ECCV) workshops*, 2018.

[14] Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., Grundmann, M., "Blazepose: On-device Real-Time Body Pose tracking," *arXiv preprint arXiv:2006.10204*, 2020.

[15] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., "Microsoft coco: Common Objects in Context," In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, pp. 740-755, 2014.

[16] Dwibedi, D., Aytar, Y., Tompson, J., Sermanet, P., Zisserman, A., "Temporal Cycle-Consistency Learning," In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 1801-1810, 2019..

KEI-CQL: A Keyword Extraction and Infilling Framework for Text to Cypher Query Language Translation

Yuan-Lin Liang, Chih-Yung Chang, Shih-Jung Wu
809416018@gms.tku.edu.tw, cychang@mail.tku.edu.tw; wushihjung@mail.tku.edu.tw;
Tamkang University, Taiwan

***Abstract*— Graph databases are invaluable for structuring and querying interconnected data, but Text-to-Cypher (CQL) remains a challenging frontier. This research paper introduces a pioneering deep learning approach that utilizes pre-trained language models to facilitate the translation of natural language queries into Cypher Query Language. Bridging the gap between these languages is crucial for organizations to manage these specific databases efficiently. This work proposes a novel framework KEI-CQL, that leverages the capabilities of deep-learning models to translate text queries into Neo4j Cypher queries. The framework KEI-CQL can extract the semantic features in the natural language query and fill in the pre-defined Cypher query sketch slots. The results from the experiment show that the proposed framework can convert the SQL query into a Cypher query.**

***Index Terms*—KEI-CQL, NL2CQL, Slot filling, BART**

I. Introduction

Knowledge graphs, hailed as powerful tools for representing structured knowledge, are pivotal in the age of information abundance. Meanwhile, semantic parsing techniques have witnessed remarkable advancements in enabling machines to understand and process human language. This paper embarks on a scholarly exploration of the union between these domains, highlighting the strategic role of semantic parsing in enhancing the construction and utilization of knowledge graph databases.

The success of pre-trained language models like OpenAI's GPT [2] series and BERT [1] from Google has sparked interest in adapting them for code-related tasks, including translation. These models, trained on vast text corpora, can capture semantic meaning, making them promising tools for understanding and generating code. Some early works [4, 5, 6] leveraged the pre-trained language models to demonstrate their utility in knowledge graph-related domains, including query translation and generation. However, most works mentioned above are performed only in a specific domain, and Cypher Query Language was not included.

This work aims to develop a framework called KEI-CQL for Text2CQL. The proposed KEI-CQL adopts the sketch-based approach, which uses multiple deep learning neural networks with a BERT pre-trained language model as an encoder to extract the entity and relation from text query and do slot filling. This work has the following contributions. First, this work represents one of the pioneering contributions to this specific task. Second, the proposed KEI-CQL developed a natural language interface for text-to-Cypher translation, which represents a user-centric contribution. Third, the proposed KEI-CQL can handle varied CQL queries.

II. Framework Architecture

Assume a standard sketch of CQL has the following format:

MATCH $Pattern ($Entity $Relationship)*
WHERE $Conn ($Entity $Op)*

Figure 1. The sketch of Cypher Query Language

The blue tokens in Fig. 1 are the slots that must be filled through the proposed KEI-CQL, and the $*Pattern* slot has its template, which is similar to the RDF template. The proposed KEI-CQL aims to translate the user query into a semantically equivalent CQL query to answer the user query without having users manually type in the CQL query. The proposed KEI-CQL consists of multiple models. The first model extracts the entity and relation from the input query and the possible attributes from the graph database. The second model is to evaluate the $*Pattern* slot wherever the filled slot has the correct pattern. The following Fig. 2 presents the details of the model.

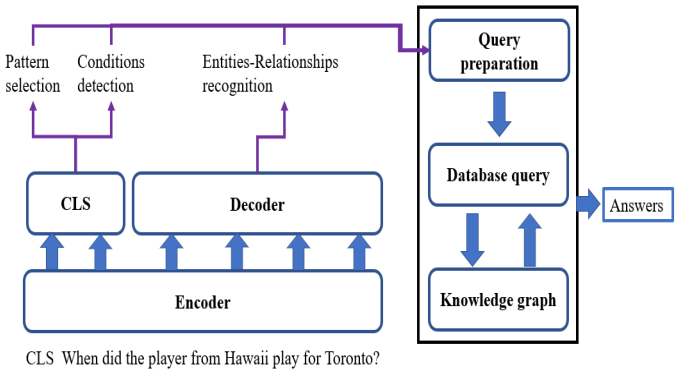


Figure 2. The KEI-CQL framework in detail.

A. Encoder model

Given an input query $Q_i = \{q_1, q_2, q_3, \dots, q_n\}$ that consists of n-th words. The input sequence is tokenized and embedded. The input embeddings will be sent into BERT model in order to obtain the semantic representation of the input sequence.

The BERT model uses weight matrices: queries Q, keys K, and values V and calculates the attention mechanism [3]. That is,

$$\text{attention}(Q, K, V) = \sum_i a_{q_i,k_i} v_i$$

After the attention mechanism, it returns the following represents a sequence of vectors:

$$H_{[\mathrm{CLS}]}, H_{q_1}, \ldots, H_{q_n}, H_{[\mathrm{SEP}]}$$

B. Output tasks

There are three tasks: pattern selection, condition detection, and entity-relationship recognition.

The pattern selection is a classification task that selects the Resource Description Framework schema (RDF) based on the input query. The task uses the sentence-level semantic representation to feed into the output layer for classification. The following equation represents the probability calculation of the pattern selection task,

$$P_{Pattern} = Softmax(W_{Pattern} H_{[\mathrm{CLS}]})$$

where $W_{Pattern} \epsilon$ R are the learnable parameters for the output probabilities $P_{Pattern}$.

Condition detection is also a classification task that determines whenever the condition is required. The task also classifies which operators will be used if the conditions are included and the relation function types if there is more than one condition. The following Equations represent the calculation of the condition detection task:

$$P_{cond} = Sigmoid(W_{cond} H_{[\mathrm{CLS}]})$$

where $W_{cond} \epsilon$ R are the learnable parameters for the output probabilities P_{cond}.

The final entity-relationship recognition task is a text generation task that extracts the entities and relationships between entities. This can capture the contextual meaning of the entities and the relationships between entities more flexibly.

III. Experiment

This research uses the spCQL [7], which is a Chinese text2CQL dataset. Each sample is a pair of Chinese Natural Language question and the equivalent CQL query. The sample also includes the answer from the executed CQL query. The pre-trained BART [8] is used as a primary encoder for the framework during training. This work is compared to the baselines that were introduced along the spCQL dataset. The results are shown in Table 1.

Table 1. The results of the logical form (lf) and execution (ex) accuracy.

Model	Validation		Test	
	Acc-lf	*Acc-ex*	*Acc-lf*	*Acc-ex*
Baseline [7]	51.3	55.2	52.8	56.2
KEI-CQL	**72.3**	**76.1**	**74.7**	**77.3**

The results demonstrate that the proposed KEI-CQL outperforms the baseline model. The ability of BART to share parameters and jointly learn multiple tasks allows it to leverage information from one task to improve performance on another. However, it is important to note that the proposed framework KEI-CQL may require more data and training time compared to the baseline.

IV. Conclusion

In conclusion, this work represents the contribution to the emerging field of text-to-Cypher transformation, an area with limited prior research. By developing and evaluating a novel approach for converting natural language queries into Cypher queries, our study addresses the gap in the domain of graph database queries. This work lays the foundation for further advancements in text-to-Cypher conversion. It has the potential to significantly enhance the usability and accessibility of graph databases for a broader audience.

References

[1] J. Devlin, M. W. Chang, and K. Lee, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol., 2019, pp. 4171–4186.

[2] A. Radford, K. Narasimhan, and T. Salimans. (2018). Improving Language Understanding by Generative Pre-Training.

[3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, "Attention is all you need", in NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, December 2017, Pages 6000–6010.

[4] S. Aghaei, E. Raad and A. Fensel, "Question Answering Over Knowledge Graphs: A Case Study in Tourism," in *IEEE Access*, vol. 10, pp. 69788-69801, 2022, doi: 10.1109/ACCESS.2022.3187178.

[5] Y. -H. Chen, E. J. -L. Lu and J. -D. Lin, "Enhancing SPARQL Query Performance With Recurrent Neural Networks," in *IEEE Access*, vol. 11, pp. 92209-92224, 2023, doi: 10.1109/ACCESS.2023.3308691.

[6] Y. -H. Chen, E. J. -L. Lu and T. -A. Ou, "Intelligent SPARQL Query Generation for Natural Language Processing Systems," in *IEEE Access*, vol. 9, pp. 158638-158650, 2021, doi: 10.1109/ACCESS.2021.3130667.

[7] A. Guo, X. Li, G.uanchen Xiao, Z. Tan, and X.Zhao. 2022. "SpCQL: A Semantic Parsing Dataset for Converting Natural Language into Cypher." In Proceedings of the 31st ACM International Conference on Information & Knowledge Management (CIKM '22). Association for Computing Machinery, New York, NY, USA, 3973–3977. https://doi.org/10.1145/3511808.3557703.

[8] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer. 2020. "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension." In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880, Online. Association for Computational Linguistics.

REsource: Energy-Efficient GPS-Based Localization for Resource-Constrained IoT Devices

Shuaibu Musa Adam, Lowie Goossens, Fedor Panafidin, Brendan J. Mackenzie, Yinze Li, Sam Michiels, Danny Hughes and Ka Lok Man

***Abstract*—In recent years, there is a growing interest in the field of battery-free and energy harvesting IoT devices. In line with this trend, we introduce BaMbI a low-cost power and communication architecture for battery-free personal communication devices. BaMbI aims to address two main challenges associated with using traditional mobile phones: i) the short lifespan of conventional phone batteries and ii) the unreliable electrical infrastructure in many parts of the world. To mitigate limited battery life, BaMbI uses super-capacitor energy storage with a projected lifetime of at least one decade. To combat unreliable charging infrastructure, BaMbI harvests energy using a solar panel and an optional hand-crank. This paper focusses on the development of an energy-efficient navigation App for use in BaMbI. Our primary goal is to significantly reduce the energy consumption of GPS-based navigation to ensure usability in BaMbI and similar devices. We build a reference hardware platform and conduct a series of experiments on the prototype, simulating different scenarios and analyzing the results. The overall results demonstrate the feasibility of our approach in decreasing the energy consumption of navigation in BaMbI.**

***Index Terms*—global positioning system (GPS), energy harvesting, resource-constrained IoT devices, navigation App, battery-free smartphone**

I. Introduction

Although most contemporary electronic devices remain battery-powered, their long-term usage is unsustainable [1], [2]. As a result, designers are exploring possible mechanisms to build electronic devices using low-power components which reduce reliance on batteries and thus impact their lifetimes and the environment. BaMbI is a battery-free and Energy Harvesting (EH) smartphone that is currently being developed at DistriNet. It offers a sustainable alternative to traditional smart phones using solar panels as the main energy source and super-capacitors energy storage. This approach presents many new challenges, as BaMbI can only store a low amount of energy compared to traditional smartphones. Therefore, to make BaMbI usable, these challenges must be addressed in both hardware and software.

While smartphone navigation is ubiquitous today, contemporary approaches depend on Global Navigation Satellite System (GNSS) technology [3] due to its worldwide precision. However, a key disadvantage of the GNSS module is its relatively high energy consumption due to long communication sessions with satellites [4], [5]. On one hand, there is also a demand for more energy-efficient localization alternatives, for example using LoRaWAN, NB-IoT or LTE-M [6]. Although these technologies show potential in energy efficiency, their availability is still limited.

EH GPS-based trackers are available for wildlife [7], [8], underwater [9], mining [10] and other applications [11]. However, to the authors' knowledge, no such system has been developed for pedestrian localization using handheld mobile devices. Against this backdrop, the primary focus of this paper is the development of battery-free and EH pedestrian navigation App for use in BaMbI. Additionally, we explore the threshold at which sensor predictions become inaccurate and the resulting implications on potential performance.

BaMbI uses super-capacitors energy storage and a solar panel together with an optional hand-crank for energy harvesting. Communication centers around LTE-M module[1], providing low-power connectivity and a Cortex-M33 processor to host application software. In terms of software, BaMbI offers intelligent charge management libraries that accurately predict and maximize device lifetime. Currently, BaMbI supports half-duplex real-time audio communication over low-bandwidth wireless networks [12], a Twitter and traffic test Apps [13]. With ongoing development and integration of more functionalities, BaMbI is evolving into a highly practical smartphone.

The remainder of this paper proceeds as follows. Section II discusses the Related Work. Section III distills the key requirements. This is followed by Section IV on sensors' Calibration, Section V describes the Implementation while Section VI focuses on Evaluation. Finally, Section VII summarizes key findings and directions for future work.

II. Related Work

GNSS signals are not always available to positioning devices, especially in dense urban areas. To improve GNSS positioning, Inertial Measurement Units (IMUs) offer a potential solution [14]–[16], as demonstrated in previous works on vehicular [14], [15], [17], pedestrian and other navigation applications. However, the most advanced IMU-GNSS algorithms for accurate positioning are proprietary and closed-source. We aim to develop an open-source IMU-GNSS localization technique that is robust to intermittent energy supply in resource-limited IoT devices.

Pedestrians face unique localization challenges when using navigation Apps, including signal loss, inaccurate route identification and long-term satellite communication needs. Ex-

Ka Lok Man is with School of Advanced Technology, XJTLU, China; and all other authors are with the imec-DistriNet, KU Leuven, B-3001 Leuven, Belgium. email: Ka.Man@xjtlu.edu.cn; firstname.lastname@kuleuven.be

[1]Long Term Evolution for Machines: LTE-M

tended usage also leads to increased energy consumption, hindering widespread adoption in battery-free IoT applications. To address these challenges, a better pedestrian localization system is needed with improved accuracy and lower energy impact compared to existing solutions.

Research has already been explored using IMU/Global Positioning System (GPS) to predict the positions of athletes when GNSS signals are temporarily unavailable [18]. Despite extensive studies on IMU-GNSS integration [15], in our opinion, there is a notable lack of focus on energy efficiency, especially in the context of battery-free and energy-harvesting platforms. This is of course, unsurprising, since GNSS module energy consumption is not a problem for battery-powered devices. Consequently, developing an energy-efficient IMU-GNSS pedestrian localization App tailored for BaMbI-like devices is crucial for minimizing power draw and ensuring optimal functionality.

III. Key Requirements

In this section, we outline key requirements for developing an energy-efficient positioning App for BaMbI. Meeting these specifications is imperative to achieving the localization goals of this paper, considering limitations of the target platform.

*a) **UI Handling**:* When navigating using a GPS-based App, the goal is always to identify and follow the desired route efficiently. Achieving this requires a fast application response to the user input [19]. REsource accomplishes this by providing a monochrome screen to display route information and draw maps, as well as a keyboard for user input and control.

*b) **Sustainable Operation**:* The application should operate solely on harvested energy and demonstrate low power consumption [8]. REsource provides a robust power management strategy, allowing reliable navigation even when solar energy is limited due to cloud cover or other adverse weather conditions.

*c) **Hardware Dependency**:* Hardware independence enables compatibility across devices and widespread adoption [20]. REsource is optimized for compatibility with BaMbI-like devices, i.e., limited memory space, persistent storage and processing power, etc. Graphical interface optimization and wireless communication compatibility also remain essential.

IV. Calibration

In REsource, we combine a GNSS module with IMU sensors to build an energy-efficient navigation App. IMU sensors require careful calibration before use in navigation Apps [21] as errors from interference and temperature changes can accumulate over time, resulting in substantial positioning drift. Through calibration, baseline readings and correction factors are established, reducing sensor noise and nonlinearity [22]. Calibrated IMUs enable accurate estimation of position over extended periods. In this section, we describe the calibrations performed on IMU sensors used in REsource, in order to minimize the aforementioned issues and optimize performance.

A. Accelerometer

The accelerometer module use in REsource is LSM6DS3[2], measuring the linear acceleration, g of the prototype in m/s^2. The Earth's gravity manifests as an acceleration vector with a typical magnitude of $9.81\ m/s^2$, varying based on location on its surface. The chosen accelerometer module also offers the option to limit the measuring range. We selected the lowest measuring range of *±2 g*, as it is sufficient for pedestrians. This module measures acceleration on 3-axes; X-axis: normal 0 m/s^2; Y-axis: normal 0 m/s^2; and Z-axis: normal -9.81 m/s^2. During measurements, it is assumed that the axes are perfectly 90^o perpendicular to each other and the prototype lies perfectly horizontal to the normal. Although these assumptions are hard to attain in practice, the imperfections are partly solved using calibration algorithm. There are 9 important factors to consider: 3 deviation factors from the linear shift of each measuring axis, 3 scale factors and 3 non-orthogonal correction factors. These factors are then applied as shown in Equation (1) [23].

$$\begin{aligned} a_p &= T_a^p \mathrm{SF}_a \left(a_m - b_a\right) \\ &= \begin{pmatrix} 1 & 0 & 0 \\ \alpha_{yx} & 1 & 0 \\ \alpha_{zx} & \alpha_{zy} & 1 \end{pmatrix} \begin{pmatrix} \mathrm{SF}_{ax} & 0 & 0 \\ 0 & \mathrm{SF}_{ay} & 0 \\ 0 & 0 & \mathrm{SF}_{az} \end{pmatrix} \\ &\times \left(\begin{pmatrix} a_{mx} \\ a_{my} \\ a_{mz} \end{pmatrix} - \begin{pmatrix} b_{ax} \\ b_{ay} \\ b_{az} \end{pmatrix} \right) \end{aligned} \tag{1}$$

Without calibration, errors persist through successive calculations, accumulating over time and producing what is known as *accelerometer drift*. Upon calibration, the speed and direction are determined according to the axis system of the object. As the results show, compensating for imperfect real-world conditions enables a reliable determination of pedestrian movement dynamics using the accelerometer as one of the key references.

B. Gyroscope

The gyroscope of the LSM6DS3, measures angular velocity of the object to which the sensor is attached in degrees per second (dps), with SI unit of rad/s. Base on this angular velocity, we can deduce what the orientation of the object is, provided that the sensor data is sufficiently reliable. To make the gyroscope data sufficiently reliable, we apply a simple calibration technique. A prototype with the gyroscope attached is placed on a flat surface, making an angle of 0^o with the horizontal. Waiting for a few seconds stabilizes the angular velocity around each axis to *zero* rad/s. As observed, the calibrated data has a variance of ±10 rad/s, enabling the gyroscope to serve as a key orientation reference for pedestrian navigation in REsource.

[2]SparkFun LSM6DS3

C. Magnetometer

REsource is equipped with a BMM150[3] magnetometer to measure magnetic fields along three perpendicular axes. This enhances precision of the altitude heading reference system filter predictions by eliminating additional measurement errors, enabling more accurate localization. Magnetometer readings can likewise be influenced by nearby magnetic materials and equipment. To address this problem, we perform a 3D calibration on the BMM150 sensor. This process involves putting the sensor in a standardized '8' motion moving along the three axes. In this way, all possible orientations of the sensor are sampled and the full range of the sensor readings captured to account for local disturbances.

D. Sensor Measurements

The sensors' measurements are read at the following frequencies:

- LSM6DS3 Accelerometer: 52 Hz
- LSM6DS3 Gyroscope: 52 Hz
- BMM150 Magnetometer: 1 Hz

The LSM6DS3 has a FIFO memory buffer, which enables the sensor to accumulate data at low power while the main application processor is in deep sleep mode. We have chosen a buffer of 200 measurements before the memory is read, allowing the controller to only read the sensor twice per second, conserving energy.

After reading, the accelerometer and gyroscope values are averaged once per second. These averages, along with magnetometer and GNSS data, are written to an external SD card. The collected data is then analyzed in MATLAB.

V. Implementation

We implemented an open-source prototype of REsource. Section V-A describes the hardware implementation while Section V-B outlines the software implementation.

A. Hardware Implementation

REsource is based around the Adafruit Feather nRF52840[4] with a Cortex-M4 processor providing 1MB of Flash and 256KB of SRAM. We have chosen this microcontroller due its low power consumption and wide range of applications in the IoT domain. We use U-blox MAX-M10S[5]as the GNSS module due to its quiescent power consumption and support for optimized power-saving modes for pedestrians. The prototype is also equipped with a monochrome screen to display information and a keyboard for user input. The sensors' measurements are stored on an external SD card. Finally, the REsource prototype shown in Figure 1 is powered using harvested solar energy that is complemented with supercapacitor energy storage for use when the solar energy is unavailable.

[3]BMM150
[4]nRF52840
[5]U-blox MAX-M10S

Fig. 1. The GPS prototype

B. Software Implementation

In this paper, we implement three different control algorithms to evaluate the performance of REsource. All algorithms follow the same fundamental principle but make tradeoffs between operation mode and power consumption. We implement the software in C++ because of its widespread use and compatibility with GNSS and IMU modules. This allows hardware integration and power optimization for our custom prototype design.

a) ***Algorithm 1: Continuous Operation Mode****:* In this algorithm, the GNSS module remains continuously ON and tracks every detectable movement of the user. The continuous operation leads to high power consumption by the module and the main controller as they constantly retrieve and process measurements. One advantage of this mode is robustness in detecting incorrect measurements. Since measurements are repeated frequently, any incorrect readings are only briefly displayed to the user before being corrected. However, the constant power draw is a major downside.

b) ***Algorithm 2: Measurement at 10 Seconds Interval****:* In this algorithm, the GNSS module performs a location measurement based on received data from the satellites. After measurement, the module is put into sleep mode for the next 9 seconds, switching off all high-energy non-essential components. After 10 seconds, the module wakes to take another location measurement after a hot start. A hot start means that the module can still use enough information from the previous positioning to quickly calculate the new location using only newly received data. For instance, the Ublox MAX-10S can become active again in just 1 second from sleep during a hot start, whereas a cold start takes an average of

28 seconds to determine the location. This hot start allows frequent sleeping to conserve energy while minimizing the time needed for data re-acquisition on each wake cycle.

*c) **Algorithm 3: Measurement at 60 Seconds Interval**:* In this algorithm, the user location is acquired using the GNSS module. After which, the module is put into sleep mode for the next 59 seconds while continuing with the measurements using the IMUs. Once the sleep time of the GNSS module elapses, the positioning is re-started again. Note that the chosen sleep time period is still within the hot start time-frame of the GNSS module. Figure 2 shows a schematic of the three combined algorithms, where continuous monitoring is considered to be *ground truth*.

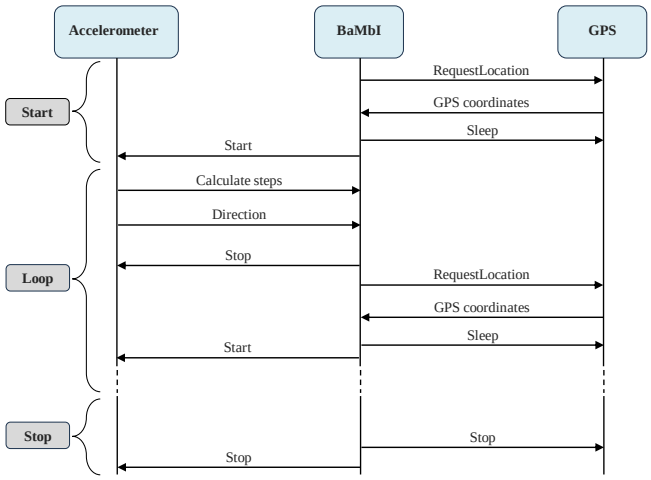


Fig. 2. Combined schematic of the three algorithms

VI. Evaluation

This section evaluates the performance of the proposed algorithms for pedestrian navigation.

A. Positioning Accuracy

The analysis of route measurement results is shown in Figure 3. The result shows a linear increase in deviation from the ground truth between pollings of the GPS. *Algorithm 3* demonstrates an approximate maximum divergence of about 50 meters, while *Algorithm 2* exhibits a reduced maximal deviation of roughly 10 meters. Therefore, this highlights the effectiveness of more frequent positioning updates in minimizing errors between recorded and actual paths.

Similarly, the route tracking accuracy of the three algorithms is shown in Figure 4. It is observed from the results that *Algorithm 1* tracks the route with a negligible error since it continuously updates the route information. *Algorithm 2* (Average Error: 1.02%, Maximum Error: 9.67%), updating every 10 seconds, also follows the route closely but has slightly more deviation when compared to the first algorithm. Whereas *Algorithm 3* (Average Error: 1.21%, Maximum Error: 9.67%), which updates only every 60 seconds, has much higher errors in tracking the route; leading to a wide gap in accuracy between the algorithms. Further, it is observed

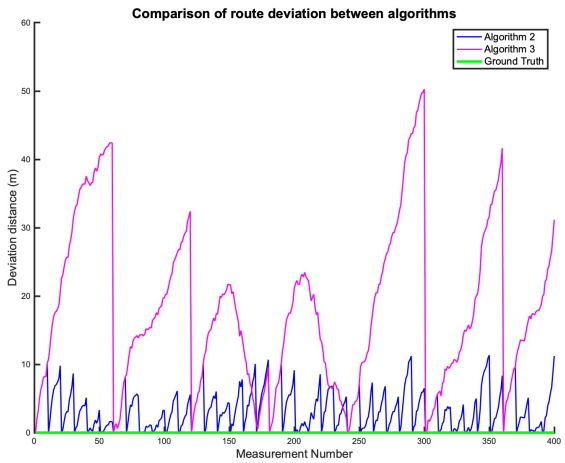


Fig. 3. A graph of measured route deviation per algorithm

that *Algorithm 3* tracks the routes more accurately at the line of site than when the route involves a fork or a turn.

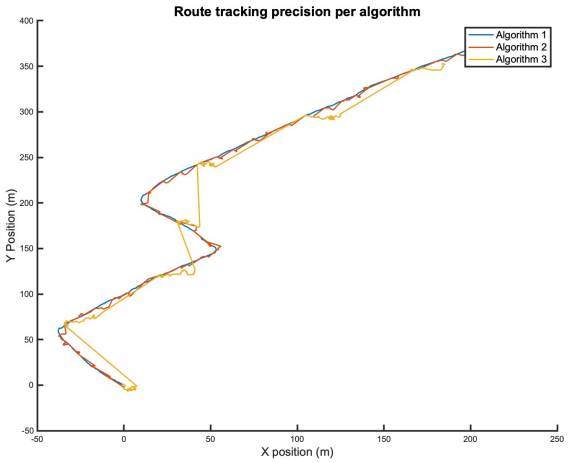


Fig. 4. Trajectory result

A potential optimization currently under exploration to balance accuracy and power could involve applying a positioning update frequency that automatically adjusts based on proximity to route changes. The conceptual approach would be that when users are close to decision points on their route, the route update temporarily increases to ensure successful navigation. Likewise, when farther from intersections or areas requiring precision, the updates could become less frequent to conserve device power.

B. Energy Consumption

Moreover, we evaluate the reliability of REsource's power consumption while running on a 5V super-capacitor energy storage. Results of the total energy draw per algorithm at 3.3V over a covered route are shown in Figure 5.

It can be seen from the figure that the energy consumption of *Algorithm 1*, the continuous operation mode is about three times higher than that of *Algorithms 2* and *3* when compared.

The substantial gap between Algorithm 1 and the other two highlights the large energy savings that intermittent localization can provide. Operating at 26% of the power required for continuous mode, *Algorithm 3* demonstrates the potential for long-term pedestrian navigation on harvested energy. However, *Algorithms 2* strikes a balance, reducing consumption to 37% of the continuous mode while maintaining sufficient route accuracy. By dynamically switching navigation mode between the three algorithms, energy-efficient localization is attainable in BaMbI.

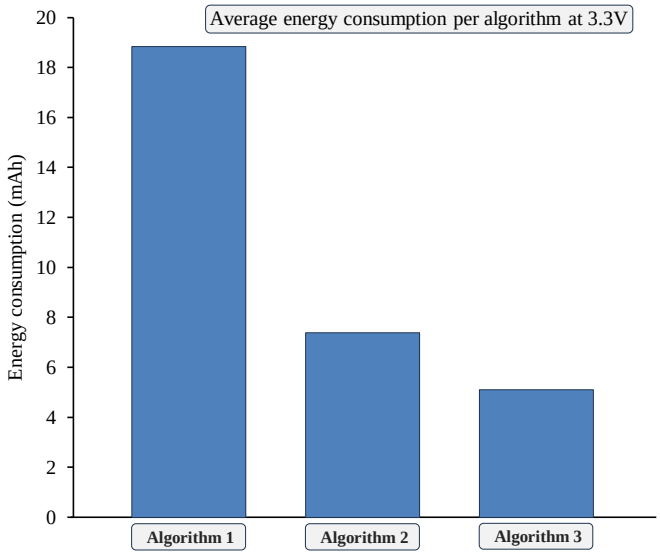


Fig. 5. Average energy consumption per algorithm at 3.3V

VII. Conclusion and Future Work

Location-based applications on resource-constrained IoT devices suffer from high energy consumption, reducing operability. To address this problem, we use IMU-GNSS fusion to reduce navigation energy usage while maintaining accuracy. We implement and evaluate a prototype for BaMbI, proposing three control algorithms to manage operations and energy consumption. Experiments analyze energy usage, accuracy and practicality. Our evaluation demonstrates an improvement in navigation power consumption using IMU-GNSS, outlining a direction for reliable positioning on BaMbI-like devices.

Finally, our future work focusses on three key areas. First, we develop a combined algorithm that dynamically switches according to route type for improved accuracy and optimal energy consumption. Second, implementing advanced pedestrian motion modeling to enable better prediction. Third, developing adaptive sampling and processing for optimized power consumption. We also plan to perform extensive real-world evaluations across diverse mobility scenarios and conditions over long time periods. By improving sensor fusion, motion models, adaptivity and field testing, we can build upon our results to realize an efficient, practical IMU-GNSS localization solution for battery-free IoT platforms.

Acknowledgements

The work presented in this paper is supported by the KU Leuven Research Fund.

References

[1] F. Yang, A. S. Thangarajan, Ramachandran, and D. Hughes, "AsTAR: Sustainable energy harvesting for the internet of things through adaptive task scheduling," *ACM Trans. Sen. Netw.*, vol. 18, no. 1, oct 2021. [Online]. Available: https://doi.org/10.1145/3467894

[2] J. de Winkel, V. Kortbeek, J. Hester, and P. Pawełczak, "Battery-free game boy: Sustainable interactive devices," *GetMobile: Mobile Comp. and Comm.*, vol. 25, no. 2, p. 22–26, sep 2021. [Online]. Available: https://doi.org/10.1145/3486880.3486888

[3] Z. Z. M. Kassas, J. Khalife, K. Shamaei, and J. Morales, "I hear, therefore I know where I am: Compensating for GNSS limitations with cellular signals," *IEEE Signal Processing Magazine*, vol. 34, no. 5, pp. 111–124, 2017.

[4] B. Karki and M. Won, "Characterizing power consumption of dual-frequency GNSS of smartphone," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.

[5] T. Janssen, A. Koppert, R. Berkvens, and M. Weyn, "A survey on IoT positioning leveraging LPWAN, GNSS, and LEO-PNT," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11 135–11 159, 2023.

[6] Y. Li, Y. Zhuang, X. Hu, Z. Gao, J. Hu, L. Chen, Z. He, L. Pei, K. Chen, M. Wang, X. Niu, R. Chen, J. Thompson, and Ghannouchi, "Toward location-enabled IoT (LE-IoT): IoT positioning techniques, error sources, and error mitigation," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4035–4062, 2021.

[7] E. Bäumker, L. Conrad, L. M. Comella, and P. Woias, "A fully featured thermal energy harvesting tracker for wildlife," *Energies*, vol. 14, no. 19, 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/19/6363

[8] T. Gregersen, T. Wild, L. Havmøller, Møller, and M. Wikelski, "A novel kinetic energy harvesting system for lifetime deployments of wildlife trackers," *PLoS ONE*, vol. 18, no. 5, p. 0285930, 2023.

[9] R. Ghaffarivardavagh, S. S. Afzal, O. Rodriguez, and F. Adib, "Underwater backscatter localization: Toward a battery-free underwater GPS," in *Proceedings of the 19th ACM Workshop on Hot Topics in Networks*, ser. HotNets '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 125–131. [Online]. Available: https://doi.org/10.1145/3422604.3425950

[10] J. P. O. W. K. Seah, R. Rayudu, and K.-L. Wang, "Accurate GPS-based ENergy harvesting Operated Relative positioning (AGENOR) for land deformation monitoring and landslide detection." 1st International Conference on Natural Hazards & Infrastructure, 2016.

[11] M. I. M. Ismail, R. A. Dziyauddin, R. Ahmad, N. Ahmad, N. A. Ahmad, and A. M. A. Hamid, "A review of energy harvesting in localization for wireless sensor node tracking," *IEEE Access*, vol. 9, pp. 60 108–60 122, 2021.

[12] S. M. Adam, Y. Liu, A.-U.-H. Ahmar, S. Michiels, and D. Hughges, "ReSoNate: A Protocol for Audio Transmission over Low Power Wide Area Networks," in *International Journal of Design, Analysis and Tools for Intergrated Circuits and Systems*, vol. 11, no. 1, 2022, pp. 6–11. [Online]. Available: https://www.cicet.org/ijdatics/issues

[13] S. M. Adam, A. S. Thangarajan, M. Liu, D. Hughes, and K. L. Man, "BaMbI, a Battery Free and Energy Harvesting Smartphone," in *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*, ser. MobiSys'22. New York, NY, USA: Association for Computing Machinery, jun 2022, p. 640–641. [Online]. Available: https://doi.org/10.1145/3498361.3538673

[14] J. Bancroft, "Multiple IMU integration for vehicular navigation," in *Proceedings of the 22nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2009)*, Savannah, GA, 2009, pp. 1828–1840.

[15] A. Angrisano, M. Petovello, and G. Pugliano, "Benefits of combined GPS/GLONASS with low-cost MEMS-IMUs for vehicular urban navigation," *Sensors*, vol. 12, no. 4, pp. 5134–5158, 2012. [Online]. Available: https://www.mdpi.com/1424-8220/12/4/5134

[16] H.-J. Chu, G.-J. Tsai, K.-W. Chiang, and T.-T. Duong, "GPS/MEMS INS data fusion and map matching in urban areas," *Sensors*, vol. 13, no. 9, pp. 11 280–11 288, 2013. [Online]. Available: https://www.mdpi.com/1424-8220/13/9/11280

[17] B. Reuper, M. Becker, and S. Leinen, " Benefits of Multi-Constellation/Multi-Frequency GNSS in a Tightly Coupled GNSS/IMU/Odometry Integration Algorithm ," *Sensors*, vol. 18, no. 9, p. 3052, sep 2018. [Online]. Available: https://doi.org/10.3390%2Fs18093052

[18] A. Waegli and J. Skaloud, "Optimization of two GPS/MEMS-IMU integration strategies with application to sports," *GPS Solutions*, vol. 13, no. 4, pp. 315–326, apr 2009. [Online]. Available: https://doi.org/10.1007%2Fs10291-009-0124-5

[19] C. Ranasinghe, S. Heitmann, A. Hamzin, M. Pfeiffer, and C. Kray, "Pedestrian navigation and GPS deteriorations: User behavior and adaptation strategies," in *Proceedings of the 30th Australian Conference on Computer-Human Interaction*, ser. OzCHI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 266–277. [Online]. Available: https://doi.org/10.1145/3292147.3292154

[20] D. Balsamo, M. Magno, K. Kubara, B. Lazarescu, and G. V. Merrett, "Energy Harvesting Meets IoT: Fuelling Adoption of Transient Computing in Embedded Systems," in *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*, 2019, pp. 413–417.

[21] G. Bakalli, A. Radi, N. El-Sheimy, R. Molinari, and S. Guerrier, "A computational multivariate-based technique for inertial sensor calibration," in *Proceedings of the 30th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2017*, Portland, Oregon, 2017, pp. 3028–3038.

[22] X. Zhu, Y. R. Zheng, and G. Dai, "Low-cost calibration of inertial measurement unit for accurate position estimation," in *Global Oceans 2020: Singapore – U.S. Gulf Coast*, 2020, pp. 1–8.

[23] M. Sipos, P. Paces, J. Rohac, and P. Novacek, "Analyses of triaxial accelerometer calibration algorithms," *IEEE Sensors Journal*, vol. 12, no. 5, pp. 1157–1165, 2012.