IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 22 September 2023, accepted 3 October 2023, date of publication 9 October 2023, date of current version 17 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3322958

RESEARCH ARTICLE

Cluster-N-Engage: A New Framework for Measuring User Engagement of Website With User Navigational Behavior

ZHOU-YI LIM, LEE-YENG ONG, (Senior Member, IEEE), AND MENG-CHEW LEOW, (Senior Member, IEEE)

Faculty of Information Science and Technology, Multimedia University, Melaka 75450, Malaysia

Corresponding author: Lee-Yeng Ong (lyong@mmu.edu.my)

This work was supported by Telekom Malaysia Research and Development under Grant RDTC/221073 (MMUE/230002).

ABSTRACT User engagement is a part of user experience that measures how attracted the users are in using a certain product or services. The level of user attention while navigating a website is a key factor in determining its effectiveness. Web usage mining can be performed on the weblog of a website to extract user navigational behavior and gain valuable insights into the useful activities of the users. The user navigational behavior can then be further evaluated to measure the user engagement of the website. This paper presents a new framework that uses user navigational behavior extracted from web usage mining to measure the engagement level of the users on the website. In this proposed framework, web session clustering is first performed on the pre-processed weblog to group similar user access patterns from web sessions. After that, four proposed engagement metrics, which are the hourly activeness, the hourly traffic, the daily activeness, and the daily traffic are calculated for each cluster to determine the engagement level of the users on the website. According to the engagement level determined by the user engagement metrics, an attention score is formulated to show how attracted the users in the session cluster are to the website. In this paper, two weblogs from different websites are used to measure their user engagement with the Cluster-N-Engage framework. The framework shows that attention score can assess the website's effectiveness in achieving its objectives.

INDEX TERMS User attention, user engagement, user navigational behavior, web session clustering, web usage mining.

I. INTRODUCTION

Over the years, the Web has grown immensely as Internet users interact with different website types. Many organizations grasp the opportunity to use websites to share and promote information about their products or services [1], [2]. In online advertising, the website effectiveness is often stringed together with user engagement. It is crucial to ensure that the users are not only engaged but also have positive feelings when they are navigating said websites [3], [4]. Capturing user's attention on a website is important for better information retrieval and processing, leading to more effective advertising on products or services [5], [6], [7]. Hence, a website should be carefully developed to ensure user's attention is retained.

User engagement refers to the quality of user experience that prioritizes how captivated and attracted a user is towards a website, and so being motivated to use it [8]. Users are engaged when they have invested their time, attention and emotion while navigating the website. Measuring user engagement of a website is not a straightforward task. In existing works, measuring user engagement commonly involves three approaches, which are self-reported engagement assessment, cognitive engagement assessment, and online behavioral metrics [8], [9]. Self-reported engagement assessment uses questionnaires or interviews to get user engagement attributes. Many researchers have developed and improved upon the existing questionnaires, such as the User

The associate editor coordinating the review of this manuscript and approving it for publication was Nazar Zaki.

Engagement Scale (UES) to measure user engagement [10], [11], [12]. The main limitation of this method is its reliance on subjectivity [13]. This is because questionnaires require users to provide their opinions and perceptions. Users may have different interpretations of the questions, which can further affect the reliability of the collected response. Cognitive engagement assessment uses follow-on task performance and physiological measures in its evaluation. The evaluation of follow-on task performance is assessed by the user's proficiency in a different task immediately after engaging in a previous task. To capture physiological data, various methods such as eye trackers, camera, and mouse pressure, can measure user engagement. While this approach is comparatively more objective than the first approach, it is more invasive and can only be used in a close quarter. Online behavioral metrics are extracted and assessed solely based on the web usage data of the website. Two common examples of the metrics that can be used are time spent on the website and click-through rates. The main downside of online behavioral metrics is that it cannot explicitly explain why users are engaged. While online behavioral metrics may not provide a straightforward way to explain user engagement, they offer a non-intrusive approach to measure it. Web usage mining (WUM) can be conducted on the website's weblog to gain insights into user navigational behavior.

WUM is a category within web mining that performs data mining techniques to extract user navigational behavior from a weblog [14], [15]. The user navigational behavior is useful in analyzing how the web pages are accessed by users, thus enhancing the website's structure and content [16]. WUM comprises three main phases, which are the pre-processing phase, the pattern discovery phase, and the pattern analysis phase, as illustrated in Fig. 1 [17], [18]. As the raw weblog does not contain useful information, the weblog is first pre-processed to be converted into a list of web sessions. Pattern discovery uses various techniques such as statistical analysis, association rule mining, and clustering to extract user behavioral patterns. In the pattern analysis phase, interesting patterns and information are extracted.

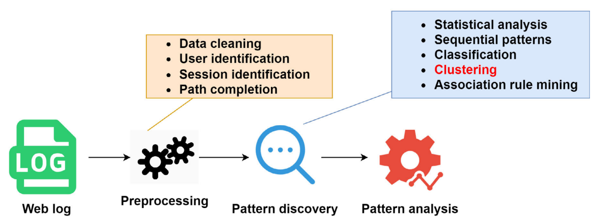

FIGURE 1. Phases of web usage mining (WUM).

To measure the user engagement of the website, user engagement metrics, also referred to as the online behavioral metrics (the third approach of user engagement assessment) can be used. Previous studies have proposed to create a single engagement score by combining multiple user engagement metrics. Aggregating multiple metrics into a single score offers several advantages that aids the decision-making process. This approach simplifies complex user navigational behavior into a concise measure, providing organizations with the overall engagement of a website. The score can be further categorized into different levels of engagement, such as high and low engagement for easier interpretation. This facilitates the comparison of the overall engagement across various websites. With the score, organizations can identify patterns and changes in the engagement of websites more efficiently. Although the engagement score in existing works is effective in evaluating the overall engagement of a website, it falls short in identifying the factors that contribute to website engagement. The contributing factors in this context refer to the user navigational behavior within the website. For organizations seeking to enhance their websites, this information is useful in providing insights into navigational behavior that drives engagement. Therefore, a new framework, known as the Cluster-N-Engage, which uses web usage mining and user engagement metrics is proposed to measure user engagement from the clusters of user navigational behavior on a website. User engagement of a website is closely related to user attention. Users invest more attention towards a website if they feel engaged and involved. The involvement level of the users correlates with the interest level and the relevance of the website [19]. Hence, an attention score is formulated to describe the intensity of user attention of a specific session clusters towards the website based on the user navigational behavior.

In the Cluster-N-Engage framework, weblogs are first pre-processed before being grouped into different-session clusters by performing web session clustering. The engagement levels of the clusters are then determined based on the four proposed user engagement metrics, namely the hourly activeness, the hourly traffic, the daily activeness, and the daily traffic According to the engagement level determined by the user engagement metrics, an attention score is calculated to show how attracted the users in the session cluster are to the website. Instead of showing the overall engagement of the website, the framework shows the user attention level of various session clusters within the website. Based on the attention scores, each session clusters are categorized into high attention cluster, medium attention cluster and low attention cluster. Organizations can explore whether session cluster that involves the target action of the website attract high user attention. User attention is measured by how active and frequent the users visit a website. Thus, if users visit a website actively and frequently, they are investing more attention to the website, leading to a good engagement. Two weblogs are used to show the applicability of this framework in measuring the user engagement for different websites. There are two main contributions in this paper, which is summarized:

- To propose a new framework that evaluates the user engagement of websites by clustering the navigational behavior from web usage mining.
- To formulate an attention score that aggregates engagement metrics that reflect the user attention level of the user navigational behavior clusters.

TABLE 1. Engagement metrics in existing studies.

References	Context	Engagement metrics	Description
Lehmann et al. [8]	User interaction in 80 websites of various types	Number of unique users [a]	Number of unique visitors
		Number of visits [c]	Number of sessions
		Number of clicks [b]	Number of page views
		Click depth [c]	Average number of page views per visit
		Dwell time per visit [d]	Average time per visit
		Active days [d]	Number of days a user visited the website
		Return rate [c]	Number of times a user visited the website
		Dwell time on website [d]	Average time a user spent on the website
Liang [20]	Mobile wallet application	Activity [b]	Total number of clicks during a session
		Duration [d]	Time spent on the website during each session
Sakas & Carrabis[21]	Airline website	Average visit duration [d]	Average time spent on the website during each session
		Pages per visit [b]	Number of pages in each visit
		Unique visitor [a]	Total number of unique visitors during a month
Bahry et al. [22]	Municipal website	Bounce rate [c]	Percentage of visits to the website that comprises a single page
		Daily page views per visitor [b]	Daily unique page views per visitor on the website
		Daily time on site [d]	Daily time spent on website per user

[a]Engagement metrics that are related to users
[b]Engagement metrics that are related to clicks
[c]Engagement metrics that are related to visits
[d]Engagement metrics that are related to time

The rest of the paper is organized as follows. Section II provides the related works in this study. Section III and IV present the description of datasets and the methodology of the framework in greater detail. The results and discussion are presented in Section V and VI, respectively. Finally, the paper is concluded with a discussion on the contribution of this work and its future work in Section VII.

II. RELATED WORKS

A. USER ENGAGEMENT METRICS

There are various studies that use different engagement metrics to measure the user engagement on a website. Table 1 shows the list of engagement metrics that are used on various websites.

Lehmann et al. [8] proposed three types of engagement metrics, which reflect popularity, activity, and loyalty. Popularity metrics, which are the number of unique users, number of visits, and number of clicks, measure how much a website is used. Activity metrics, which are the click depth and dwell time per visit, measure how a website is used by the users. Loyalty metrics, which are the active days, return rate, and dwell time on the website, measure how frequently users return to a website. The website is said to be highly engaging if the three types of engagement metrics are high. Liang [20] selected the number of clicks per visit and time spent per visit as user engagement metrics to measure the relationship between user engagement and purchase decision in mobile wallet platform. In [21], user engagement metrics such as average visit duration, pages per visit, and number of unique visitors are used to illustrate the correlation between the website traffic and the user engagement. Bahry et al. [22] used bounce rate, daily page views per visitor, and daily time on site to measure user engagement of the municipal websites.

In summary, various user engagement metrics are used in existing studies, depending on the specific type of engagement that the authors aim to investigate. However, they can be broadly categorized into the metrics based on the users, the clicks, the website visits, and the time. They are represented by *a*, *b*, *c*, and *d* respectively in Table 1. Each user engagement metrics measures the different aspects of a website. As a single metric is insufficient to evaluate the engagement of a website, many studies use multiple metrics for this purpose. However, having multiple metrics make it challenging to make decisions as they might provide conflicting insights into user behavior. These works do not aggregate the various metrics into a single score that can conclude the engagement of a website. Having an engagement score is useful in assessing the users' overall impression of a website, which allows organizations to identify areas where the website can be improved. This will be discussed in the next subsection.

B. USER ENGAGEMENT ASSESSMENT

In the previous subsection, user engagement metrics show only the engagement of different aspects of the website. However, having an aggregated engagement score is important to show the engagement level of a website with the aid of the calculated metrics. This subsection presents three studies that perform user engagement assessment. Table 2 depicts the engagement assessment in the existing studies.

Peterson and Carrabis [23] defined visitor engagement as the measurement of the depth of user interaction against a clearly defined set of goals. The authors then proposed a new calculation method for measuring visitor engagement. This was done by calculating the click depth index, duration index, recency index, brand index, feedback index, interaction index, and loyalty index. Each of these metrics

TABLE 2. Engagement assessments in existing studies.

References	Engagement calculation	Engagement level
Peterson & Carrabis [23]	$$E = (C_i + D_i + R_i + L_i + B_i + F_i + I_i)$$ where E is the engagement score, C_i=Sessions which have at least 4 page views/All sessions, D_i=Sessions where duration is over 5 minutes/All sessions, R_i=1/Number of days elapsed since the most recent session, L_i=1-(1/Number of sessions during the timeframe), B_i=Sessions where it contains user searched a phrase related to the brand/All sessions, F_i=Number of positive responses/All qualitative questions asked, I_i=Sessions where users complete an action/All sessions	No distinct levels of engagement
Lagun & Lalmas [24]	$$E = (\frac{Bounce}{N}, \frac{Shallow}{N}, \frac{Deep}{N}, \frac{Complete}{N})$$ where E is the engagement score, N=Total number of page views for the article, Bounce=Number of page views that are categorized as Bounce, Shallow=Number of page views that are categorized as Shallow engagement, Deep=Number of page views that are categorized as Deep engagement, Complete=Number of page views that are categorized as Complete engagement	**Bounce**: Users who do not engage with the article and leave the page within 10 seconds **Shallow**: Users who decide to stay and read the article but reads less than 50% of it **Deep**: Users who decide to stay and read the article but reads more than 50% of it **Complete**: Users who finish reading the article and decide to post or reply to comments
Fedushko et al. [25]	$$U_i = \frac{\max(R_i, C)}{\max(V_i, C)} W_i$$ where U_i is the engagement score, R_i=Number of reactions on a certain online project topic, V_i=Number of views or visitors or subscriptions on a certain online project topic, C=A constant used to avoid getting 0 during division, W_i=A weight parameter which is different for each topic and its U_i	**Low engagement**: Score between 0.00 and 0.05 **Moderate engagement**: Score between 0.05 and 0.65 **High engagement**: Score between 0.65 and 1.00

aims to measure engagement from various parts of the website. A high engagement score shows the users are highly engaged on the website. While this method can yield a single score to show the overall engagement of the website, the authors did not provide any additional information regarding classifying the score into distinct levels of engagement. Knowing the level of engagement of a website is a straightforward way to understand the interest level of the users.

In [24], the authors proposed four engagement metrics that are based on variations of the dwell time to measure the user engagement in online news reading. Online news website comprises many news articles of different topics. The engagement metrics used are the dwell time on page, the dwell time on page header, the dwell time on body of page, and the time spent on the comment section of page. The authors first performed a clustering algorithm on the dataset before measuring each cluster with the engagement metrics to identify the different page reading patterns. A set of engagement levels are suggested by the authors to show user attention on a news article, which include bounce, shallow engagement, deep engagement, and complete engagement. To calculate the article user engagement profile, a probabilistic model is formulated. This probabilistic model is effective in showing the proportion of each engagement level in the article. However, this engagement assessment method is not applicable in different types of websites as the work focus solely on online news website. In addition, this approach measures the article's engagement solely based on the percentage of the content read. This may not be suitable for other website types that demand additional metrics to evaluate engagement.

In a study by Fedushko et al. [25], the authors proposed a method for technical condition assessment by combining both user engagement and Service Level Indicators (SLIs)/Service Level Objectives (SLOs)/Service Level Agreements (SLAs) measurements to identify types of user satisfaction with infrastructure state of online learning services such as Coursera and Udemy. The authors presented a method for calculating the engagement score of the website. After the engagement score is obtained, it is standardized to the range from 0 to 1. Based on the standardized engagement score, the engagement level of the website is categorized into low, moderate, or high. The engagement levels are mapped with the SLO/SLA to show how websites can be improved upon. Although the focus of this study is to measure the user satisfaction of online learning services, this method shows how having an engagement score and distinct levels of engagement can help understand user behavior.

In summary, each user engagement assessment method in previous studies have their own strengths and weaknesses.

TABLE 3. Web page categories-online roadshow website.

Page name	URI keywords	Page name	URI keywords
pre-survey page	pre_survey_page	FAQ page	faq
landing page	landing_page	game 1	Flappy Bird
login page	login_page	game 2	Brick Breaker
dashboard page	dashboard_page	game 3	E-motion
redemption history page	redeem_history_page	game 4	Find the Invisible Dog
redemption page	shop_page	game 5	Voice Control
registration page	register_page	game 6	Pose Matching
logout page	logout		

The proposed Cluster-N-Engage framework aims to overcome the limitations of previous studies. The similarities and differences between Cluster-N-Engage framework and the others are highlighted:

- In [24] and [25], these frameworks are applicable only to particular types of websites, which are online learning services and online news website, respectively. Like [23], the Cluster-N-Engage framework is designed to work on different websites.
- Like [24], the Cluster-N-Engage framework measures engagement by solely utilizing the user navigational behavior extracted from weblogs. The engagement metric used to calculate the engagement score in [24] is the number of page views. The weblog alone does not measure the engagement metrics listed in [23] and [25], as they require external sources such as questionnaires.
- Existing frameworks in [23], [24], and [25] produced a single engagement score that shows the overall engagement of the website. The Cluster-N-Engage framework produces a single engagement score as well, which will be referred to as the attention score. The attention score shows the level of user attention directed to the website. Both the engagement score and proposed attention score combine multiple engagement metrics into a single measure, which can simplify decision-making process.
- In [23], the final engagement score was not classified into distinct levels of engagement. In contrast, the Cluster-N-Engage framework will classify the attention score into one of three levels: high attention, medium attention, and low attention. Categorizing the score into engagement levels provides a clearer interpretation of user behavior within the website.
- The primary difference between the Cluster-N-Engage framework and previous works lies in the user engagement assessment of the website. Engagement score in previous works only provides the overall engagement of the website. The Cluster-N-Engage framework addresses this issue by performing clustering and introducing attention score. Once the user navigational patterns are first clustered based on their similarity, attention score is used to measure the level of user attention for each cluster. This method can then identify the user attention level of the clustered navigational behavior within the website.

III. DATASET DESCRIPTION

This section briefly describes the weblogs of different sizes used in this study. Two weblogs are used in this study to test the applicability of the framework on different datasets. The first weblog is obtained from a ready online roadshow website [26], [27]. Online roadshow is a roadshow in the virtual environment that promotes products or services on digital devices. The main aim of the online roadshow website is to play games to redeem vouchers. The products-specific advertising elements are embedded into the games to engage the users in the online roadshow website. In this ready online roadshow website, new users are required to fill up a survey regarding the campaign, which will lead them to the landing page. The user is then prompted to create an account to join the campaign. After that, the user is directed to the dashboard page, which displays six available games. The six available games contain information that the campaign intends to convey. The user must complete all six games to redeem vouchers at the redemption page. When the user logs out, they will be led to the logout page.

The online roadshow website contains 20,135 entries in its weblog, which are collected from July 5, 2021, to August 31, 2021, spanning 57 days. There are 15 pages after the weblog is pre-processed. Within these pages, this study represents the six game pages as game 1, game 2, game 3, game 4, game 5, and game 6. Table 3 presents the description of web pages of the online roadshow website.

The second weblog is a dataset from an e-commerce website taken from Harvard Dataverse [28], [29]. The concept of the website is elaborated based on existing e-commerce websites. An e-commerce website is a website that allows users to purchase items or services on their digital devices. The main aim of the e-commerce website is to purchase products. To use the website, users are first required to register for an account. Users can search for products to purchase by browsing a list of products and using the product search bar. When users have the intention to purchase a product, they can add the product into the shopping cart. They can place an order for the product and make a payment if they intend to purchase it. Users can also recommend a product to their peers. After the users receive the products, they can leave a product review on the website.

The weblog records user visits to the e-commerce website for a duration of 30 days, specifically from April 1, 2020, to April 30, 2020. It comprises 5,597,529 entries. After

TABLE 4. Web page categories-E-commerce website.

Page category	URI keywords	Page category	URI keywords
product page	• product_ • product_inquiry • products_browse • product_details • product • product_edit • product_list	delivery page	• delivery_information • address_data • delivery_address • change_delivery_location • shipment_change • shipment_edit • choose_delivery_location
shopping cart page	• shopping_cart • shopping_cart_details • shopping_points • to_shopping_cart_quantity • to_shopping_cart • remove_from_cart • calculate_shopping_cart • delete_shopping_cart • save_shopping_cart • remove_shopping_cart_product • load_shopping_cart • remove_discount_coupon • active_discount_coupon • check_coupon • add_coupon	order page	• orders • order_login • order_confirmation • order_summary • order_invoice • change_order_data • order_browse • order_details • order_list_status • orders_shipping_inpost_action • orders_shipping_inpost • order_invoice • order_fulfillment • to_order_registration • to_change_order_data • to_login_order • add_order_product • order_product_discount
customer panel page	• customer_panel • password_reminder • password_change • availability_notification • change_country • to_change_password • change_password • password_reset	about page	• our-internet-shop-about-us • using_our_store • shop_rules • information_online
recommend page	• recommend • recommended • recommend_friend_product	complaint page	• write_complaint • complaints_browse • complaint_details
payment page	• end_payment • payment_change • change_purchase_document	login page	• login • to_login
category page	• category_ • list_categories	manufacturer page	• manufacturer • manufacturers
registration page	• registration • to_registration	review page	• write_review • reviews
logout page	• log_out	news page	• news
search page	• search	sale page	• sale
clipboard page	• clipboard	contact page	• contact
policy page	• privacy_policy	update page	• update
index page	• index.php		

pre-processing, the recorded URIs are further classified into 23 distinct page categories. Table 4 presents the description of the page categories of the e-commerce website.

IV. METHODOLOGY

In this section, Cluster-N-Engage framework that combines both web usage mining (WUM) and user engagement is presented. The three main phases of WUM that are illustrated in Fig. 1 are adapted into the framework to show the relationship between WUM and user engagement. Fig. 2 depicts the general overview of the framework. In the pre-processing phase, the datasets undergo data cleaning, user identification, session identification and path completion. In the pattern discovery phase, clustering, specifically web session clustering is conducted to extract session clusters present in the datasets. Finally, user engagement of the session clusters is evaluated in the pattern analysis phase. Fig. 3 provides the expanded view of the Cluster-N-Engage framework from Fig. 2. The subsequent section is organized according to the three main phases of WUM presented in Fig. 3.

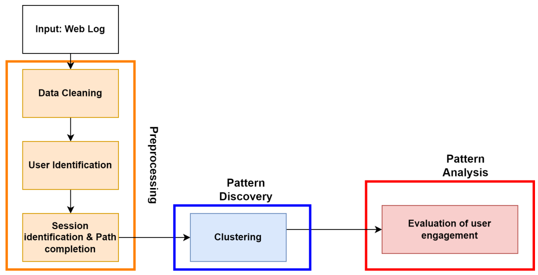

FIGURE 2. General overview of cluster-N-Engage.

A. PRE-PROCESSING PHASE

Both weblogs are first cleaned to remove irrelevant entries and to perform necessary data type conversion. The data cleaning process of both weblogs differs but they follow the common processes that are illustrated in Fig. 3. The snippets of the online roadshow weblog and e-commerce weblog are shown in Fig. 4 and Fig. 5, respectively.

The online roadshow weblog comprises nine fields, including the log ID, page, user ID, temp ID, enter page time, log time, device used to access the website, campaign ID, and IP address. Only three fields, which are the user ID, log time, and page are selected in this study as they are relevant for the clustering of web sessions based on user navigational behavior. Selecting only the essential fields helps to improve the clustering quality. Log entries that have missing records in any fields are removed from the dataset. The e-commerce weblog contains 10 fields, which are the IP ID, user ID, timestamp, HTTP method, URI, HTTP version, status code, bytes, referrer, and user agent. Based on the suggestion by [30], log entries that do not provide meaningful information during the mining process should be first removed. According to [31], log entries that do not contain the HTTP method value of GET or POST can be removed. Log entries with status codes between 200 to 299 that give a successful response are kept as they show the successful actions made by the users. In the e-commerce weblog, records of several log entries that the users did not explicitly request, such as graphic files and scripts are eliminated. The data within these files is not relevant to actual user navigational behavior. According to Yu et al. [32], log entries created by network robots can be identified by their IP address and agents. Therefore, log entries with the user agent value of 'Go-http-client', 'python', 'bot', and 'Bot' created by network robots are removed. Only four fields, which are the IP ID, timestamp, URI, and user agent are kept in this study. The names of the URIs in the e-commerce weblog are translated from Polish to English using Google Translate for easier interpretation. As the dataset contains many URIs, they are manually grouped into 23-page categories based on their relatedness, as outlined in Table 4. URIs that do not represent meaningful pages or cannot be categorized into the defined page categories are removed. The URI is represented as not meaningful if it does not provide information on user navigational behavior on the e-commerce website. Some examples of these types of URI are '/vpn/', '/other/bookmarks.php', and '/other/captcha.php'.

After data cleaning is performed on the datasets, user identification is conducted to reveal the identity of each user. For the online roadshow weblog, the users can be identified straightaway by referring to the user ID field. Log entries with the user IDs of the website administration team are removed. In the e-commerce weblog, individual users are identified based on the IP ID and user agent fields. In this proposed framework, the same user is identified from the same IP ID and user agent as suggested by Selvakumar et al. [33].

A web session contains all the actions performed by a user from the moment they enter a website until the exit. The main objective of session identification is to split the page access of each user into individual sessions, each representing a website visit as suggested by Asadianfam and Mohammadi [34]. Before session identification is conducted, all the web pages in both datasets are renamed as p1, p2, p3, and so on. This renaming process is for the vectorization process in the latter steps. The time on page is calculated by subtracting *log time of the current page* from the *log time of the next page*. The time-out technique is used in this study to detect web sessions. According to Guerbas et al. [35], the web session is identified during a pre-defined threshold time-out interval in this technique. Session identification is performed with a time-out value of 30 minutes based on a few conditions. A user is considered having exited the website, or started a new web session if any of the following conditions are met: (1) the user has accessed a page for over 30 minutes, (2) the page accessed is the last recorded page on the website, and (3) the page accessed is a logout page. Table 5 shows the description of both datasets after the sessions are identified.

TABLE 5. Data description.

Dataset	Online roadshow website	E-commerce website
Time period	5/7/2021 to 31/8/2021	1/4/2021 to 30/4/2021
Initial log entries	20,135	5,597,529
Log entries after cleaning	12,076	995,214
No of users	397	29,098
No of sessions identified	2,361	104,875

Every web session is a path sequence that is taken by the user. Pages within the same session are converted into a list, representing a path. Fig. 6 visually presents a sample list of sessions in path form. Taking the last row as an example, the path of user 762 in the session is *[p6, p5, p5, p11, p5]*, indicating that user 762 visits p6, p5, p11, and finally p5 again during the session. Both datasets are then vectorized to be converted into a large binary matrix. The matrix size is determined by the number of sessions and the number of pages. In this study, the matrix of the online roadshow weblog is $2,361 \times 15$, and the matrix of the e-commerce weblog

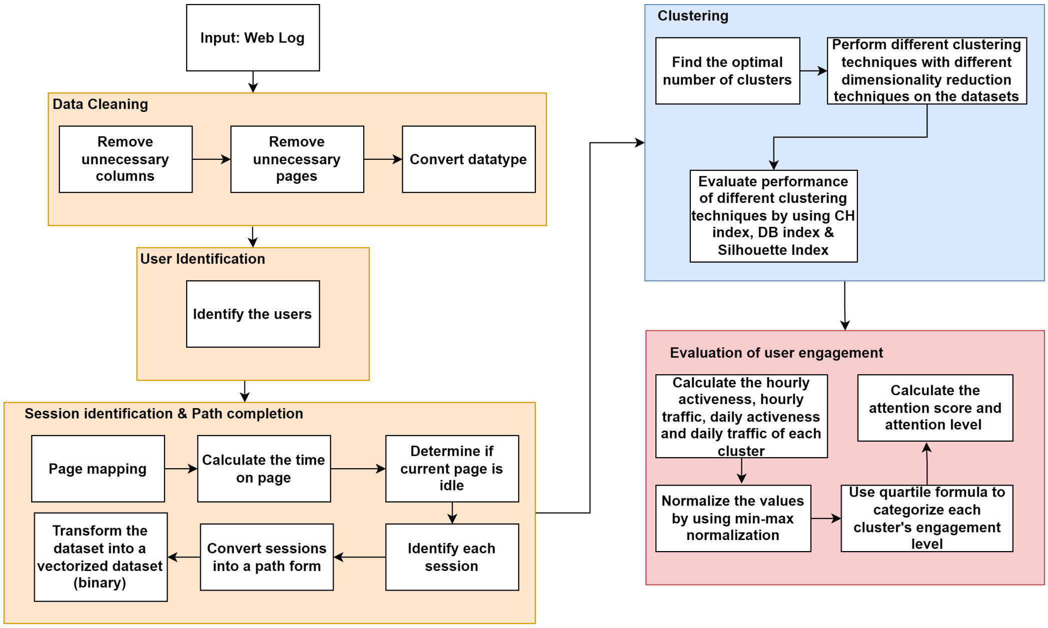

FIGURE 3. Expanded view of cluster-N-Engage.

	log_id	page	user_id	temp_id	enter_page_time	log_time	device	campaign_id	ip_address
0	35172	pre_survey_page	NaN	408.0	NaN	5/7/2021 1:25:40	phone	182.0	178.128.48.225
1	35173	pre_survey_page	NaN	NaN	NaN	5/7/2021 3:01:32	phone	182.0	178.128.48.225
2	35174	pre_survey_page	NaN	NaN	NaN	5/7/2021 4:59:49	phone	182.0	178.128.48.225
3	35175	pre_survey_page	NaN	NaN	NaN	5/7/2021 5:00:36	phone	182.0	178.128.48.225
4	35176	pre_survey_page	NaN	NaN	NaN	5/7/2021 6:04:36	phone	182.0	178.128.48.225

FIGURE 4. Snippet of online roadshow weblog before pre-processing.

IpId	UserId	TimeStamp	HttpMethod	Uri	HttpVersion	ResponseCode	Bytes	Referrer	UserAgent
1PL	-	637107516000000000	GET	/p-9710.html	HTTP/1.1	200	14226	https://shop.our-internet-company.pl/p-3880.ht...	Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.3...
1PL	-	637107516000000000	GET	/szablony/shop_11.rwd/css/style.php?ncss=style...	HTTP/1.1	200	19214	https://shop.our-internet-company.pl/p-9710.html	Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.3...
1PL	-	637107516000000000	GET	/javascript/skrypty.php	HTTP/1.1	200	29403	https://shop.our-internet-company.pl/p-9710.html	Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.3...
1PL	-	637107516000000000	GET	/javascript/produkt.php	HTTP/1.1	200	7914	https://shop.our-internet-company.pl/p-9710.html	Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.3...
1PL	-	637107516010000000	GET	/szablony/shop_11.rwd/obrazki/szablon/info.png	HTTP/1.1	200	464	https://shop.our-internet-company.pl/szablony/...	Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.3...

FIGURE 5. Snippet of e-commerce weblog before pre-processing.

user_id	path
762	[p1, p2, p3]
762	[p3]
762	[p2, p3, p3, p4, p5, p7]
762	[p6, p5, p5, p11, p5]

FIGURE 6. An example of a list of sessions in path form.

is 104, 875 $\times$ 23. Each cell within the matrix shows if a particular page has been requested during a specific session, represented by 1 for yes and 0 for no. Table 6 shows the paths from Fig. 6 in a matrix form as an example.

B. PATTERN DISCOVERY PHASE

To analyze the user navigational behavior in the datasets, web session clustering is performed on the matrices generated during the pre-processing phase. Since the initial number of clusters of both datasets is unknown, the optimal number of clusters is determined by using the Elbow Method with the Sum of Squared Error (SSE) [36]. The optimal number of clusters is selected by finding the "elbow point".

TABLE 6. An example of a binary session matrix.

p 0	p 1	p 2	p 3	p 4	p 5	p 6	p 7	p 8	p 9	p1 0	p11
0	1	1	1	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0
0	0	1	1	1	1	0	1	0	0	0	0
0	0	0	0	0	1	1	0	0	0	0	1

After obtaining the optimal number of clusters, four clustering techniques, namely k-means, k-medoids, bisecting k-means, and fuzzy c-means, are performed on both web session matrices. Given the multidimensional nature of the matrices, three dimensionality reduction techniques, such as the principal component analysis (PCA), the t-distributed stochastic neighbor embedding (TSNE), and the uniform manifold approximation and projection (UMAP) are conducted on the web session matrices prior to the clustering step. To assess if dimensionality reduction enhances clustering quality, multiple setups are examined. Each clustering technique will be performed on the matrices, with and without dimensionality reduction.

The Euclidean distance is used as the similarity measure in the clustering experiments, as it is commonly applied in web usage mining [18], [37]. The clustering results are then evaluated based on the three cluster validity indices, which are the Silhouette index (SI), the Davies-Bouldin (DBI) index, and the Calinski-Harabasz (CHI) index [38].

C. PATTERN ANALYSIS PHASE

In the pattern analysis phase, the user behavior of the session clusters that are obtained from the pattern discovery phase is extracted by using the user engagement metrics. At the end of this phase, an attention score is calculated to show the level of user attention towards the website.

User engagement in the Cluster-N-Engage framework relates to how involved the users are on the website. When users are engaged, they are more likely to pay attention to a website. Four user engagement (UE) metrics, namely the hourly activeness, the hourly traffic, the daily activeness, and the daily traffic are proposed in the Cluster-N-Engage framework. The proposed metrics are derived from the two commonly used user engagement metrics, which are the number of clicks and the number of sessions. Number of clicks and number of sessions serve as indicators of the level of user attention towards the website. When users are engaged with a website, they are likely to have higher levels of navigation activity and visit the website more frequently. Number of clicks is related to the activeness of the users on the website, meanwhile the number of sessions is related to the traffic of the website. Both the number of clicks and number of sessions are calculated based on time slots and dates. The time slots are represented in the 24-hour time format, ranging from 00:00 to 23:59. The dates correspond to the time period of the datasets, is stated in Table 5. Therefore, hourly activeness and hourly traffic refer to the number of clicks and sessions made by users hourly regardless of the date. Meanwhile, daily activeness and daily traffic refer to the number of clicks and sessions made by users daily regardless of the time. The number of clicks and sessions are calculated for each session cluster, taking into consideration on the time slots and dates. Table 7 depicts the definition of the proposed UE metrics.

TABLE 7. Definition of the proposed user engagement metrics.

User engagement metrics	Description	Calculation
Hourly activeness	How active the users are in the session cluster generally at hourly intervals?	Number of clicks with respect to time slot
Hourly traffic	How high of website traffic is to the users in the session cluster generally at hourly intervals?	Number of sessions with respect to time slot
Daily activeness	How interested the users are in the session cluster generally within a certain timeframe?	Number of clicks with respect to date
Daily traffic	How popular the website is to the users in the session cluster generally within a certain timeframe?	Number of sessions with respect to date

After the UE metrics of each time slot and date of all clusters are recorded, the calculated values are subjected to max-min normalization to scale them within 0 to 1. The engagement levels of each normalized value are then determined using the quartile formula. The minimum value, maximum value, first quartile, median value, and third quartile are calculated separately for every UE metric. Table 8 shows the calculation of the quartile values for the hourly activeness metric and hourly traffic metric as an example.

The quartile formula is used to categorize the normalized values into four equal parts: *very low engagement*, *low engagement*, *medium engagement*, and *high engagement*. Fig. 7 and Table 9 illustrate how the normalized values are categorized into different engagement levels. For example,

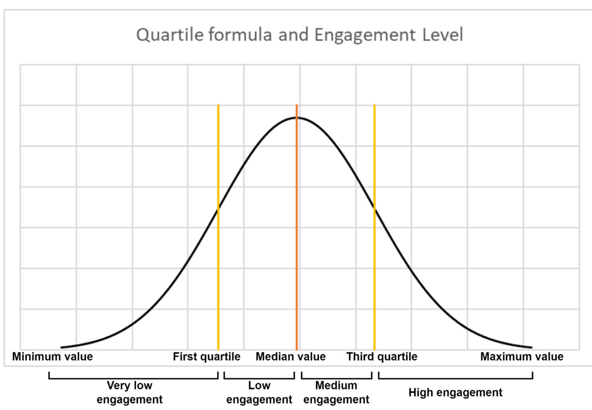

FIGURE 7. Illustration of engagement levels according to the quartiles.

			Hourly activeness			Hourly traffic
	No of clicks	Normalized no of clicks	Engagement level (clicks)	No of sessions	Normalized no of sessions	Engagement level (sessions)
00:00-00:59	73	0.145129225	High	20	0.454545455	High
01:00-01:59	26	0.051689861	Low	6	0.136363636	Low
02:00-02:59	7	0.013916501	Low	4	0.090909091	Low
03:00-03:59	0	0	Very low	0	0	Very low
04:00-04:59	0	0	Very low	0	0	Very low
05:00-05:59	5	0.009940358	Low	2	0.045454545	Very low
06:00-06:59	0	0	Very low	0	0	Very low
07:00-07:59	6	0.011928429	Low	2	0.045454545	Very low
08:00-08:59	31	0.061630219	Low	8	0.181818182	Low
09:00-09:59	28	0.055666004	Low	12	0.272727273	Medium
10:00-10:59	53	0.105367793	Medium	23	0.522727273	High
11:00-11:59	108	0.21471173	High	26	0.590909091	High
12:00-12:59	68	0.135188867	High	26	0.590909091	High
13:00-13:59	122	0.242544732	High	44	1	High
14:00-14:59	115	0.228628231	High	29	0.659090909	High
15:00-15:59	106	0.210735586	High	24	0.545454545	High
16:00-16:59	152	0.302186879	High	44	1	High
17:00-17:59	139	0.276341948	High	43	0.977272727	High
18:00-18:59	64	0.127236581	High	23	0.522727273	High
19:00-19:59	72	0.143141153	High	30	0.681818182	High

FIGURE 8. Snippet of the categorization of engagement level of the hourly activeness and hourly traffic metrics of a cluster.

TABLE 8. Snippet of the quartile values of the hourly activeness and hourly traffic metrics.

Number of clicks with respect to time slot (Hourly activeness)		Number of sessions with respect to time slot (Hourly traffic)	
Minimum value	0	Minimum value	0
First quartile	0.01193	First quartile	0.06250
Median value	0.06362	Median value	0.18182
Third quartile	0.11978	Third quartile	0.34091
Maximum value	1	Maximum value	1

TABLE 9. Engagement levels according to the quartile formula.

Quartiles	Engagement level
Minimum value to first quartile	Very low
First quartile to median value	Low
Median value to third quartile	Medium
Third quartile to maximum value	High

according to Table 8, if the normalized value of the hourly activeness metric is between the median value (0.06362) and third quartile (0.11978), it will be categorized as *medium engagement*.

Once the engagement levels of the normalized values are obtained, the overall engagement level of each session cluster based on the UE metrics can be assessed. This can be achieved by summing up the frequency of *very low engagement*, *low engagement*, *medium engagement*, and *high engagement* of each UE metric. As illustrated in Fig. 8, at the hourly traffic metric column, the frequency of *high engagement* is the highest. Therefore, the engagement level of the session cluster based on the mentioned UE metric is high.

Each UE metric shows different types of engagement that can be seen in each session cluster. Hourly activeness and hourly traffic depict a general overview of user engagement in the session cluster. Meanwhile, daily activeness and daily traffic reveal a more specific overview of user engagement of the session cluster as it takes into consideration the engagement level on a daily basis. Table 10 shows the characteristics of each cluster according to the UE metrics and its engagement level.

An attention score is introduced to describe the intensity of the user attention of a specific session cluster toward the website based on the UE metrics. The score for each UE metric depends on the engagement level: 1 for *very low engagement*, 2 for *low engagement*, 3 for *medium engagement*, and 4 for *high engagement*. The scores of the UE metrics are first summed up and then averaged, resulting in the attention score. Based on the attention score, it is then further categorized into three attention levels, as shown in Table 11. The higher attention score shows that a user visits a website actively and frequently.

V. RESULTS

This section first briefly presents the results of the clustering experiment. The results of this experimental study were reported in [39]. This section will further discuss the user engagement results by using the proposed framework in this study.

Before performing different clustering techniques on the datasets, the optimal number of clusters was identified as 10 for the online roadshow weblog and 5 for the e-commerce weblog. The clustering techniques with the best results, as determined by the three cluster validity indices (Silhouette index, DB index, and CH index), are selected for further user engagement analysis. Based on the experimental study in [39], k-medoids clustering with no dimensionality

TABLE 10. Engagement level and descriptions.

User engagement metrics	Engagement level	Cluster type
Hourly activeness	Very low	• The number of clicks is generally very low at hourly intervals. • Very low or no hourly activeness on the website. • Users of this session cluster generally have very low interaction with the website at hourly intervals.
	Low	• The number of clicks is generally low at hourly intervals. • Low hourly activeness on the website. • Users of this session cluster generally have low interaction with the website at hourly intervals.
	Medium	• The number of clicks is generally moderate at hourly intervals. • Moderate hourly activeness on the website. • Users of this session cluster generally have moderate interaction with the website at hourly intervals.
	High	• The number of clicks is generally high at hourly intervals. • High hourly activeness on the website. • Users of this session cluster generally have high interaction with the website at hourly intervals.
Hourly traffic	Very low	• The number of sessions is generally very low at hourly intervals. • Very low hourly traffic on the website. • Users of this session cluster generally seldom visit the website at hourly intervals.
	Low	• The number of sessions is generally low at hourly intervals. • Low hourly traffic on the website. • Users of this session cluster generally visit the website once in a while at hourly intervals.
	Medium	• The number of sessions is generally moderate at hourly intervals. • Moderate hourly traffic on the website. • Users of this session cluster generally sometimes visit the website at hourly intervals.
	High	• The number of sessions is generally high at hourly intervals. • High hourly traffic on the website. • Users of this session cluster generally always visit the website at hourly intervals.
Daily activeness	Very low	• The number of clicks is generally very low on a daily basis. • Very low or no daily activeness on the website. • Users of this session cluster generally have very low interaction with the website on a daily basis.
	Low	• The number of clicks is generally low on a daily basis. • Low daily activeness on the website. • Users of this session cluster generally have low interaction with the website on a daily basis.
	Medium	• The number of clicks is generally moderate on a daily basis. • Moderate daily activeness on the website. • Users of this session cluster generally have moderate interaction with the website on a daily basis.
	High	• The number of clicks is generally high on a daily basis. • High daily activeness on the website. • Users of this session cluster generally have high interaction with the website on a daily basis.
Daily traffic	Very low	• The number of sessions is generally very low on a daily basis. • Very low hourly traffic on the website. • Users of this session cluster generally seldom visit the website on a daily basis.
	Low	• The number of sessions is generally low on a daily basis. • Low hourly traffic on the website. • Users of this session cluster generally visit the website once in a while on a daily basis.
	Medium	• The number of sessions is generally moderate on a daily basis. • Moderate hourly traffic on the website. • Users of this session cluster generally sometimes visit the website on a daily basis.
	High	• The number of sessions is generally high on a daily basis. • High hourly traffic on the website. • Users of this session cluster generally always visit the website on a daily basis.

reduction achieved the best performance across all three metrics for the online roadshow website. As for the e-commerce weblog, the combination of k-medoids and UMAP is the most promising model. These two models are used in the subsequent pattern analysis phase to investigate user engagement.

TABLE 11. **Attention scoring system.**

Attention level	Score range	Description
High	Score between 3.00 to 4.00	• The website receives a high volume of user attention in this session cluster. • The website receives high user involvement in this session cluster. • Users of this session cluster have highly invested their time and attention on this website.
Medium	Score between 2.00 to 2.99	• The website receives a moderate volume of user attention in this session cluster. • The website receives moderate user involvement in this session cluster. • Users of this session cluster have moderately invested their time and attention on this website.
Low	Score between 1.00 to 1.99	• The website receives a low volume of user attention in this session cluster. • The website receives low user involvement in this session cluster. • Users of this session cluster have slightly invested their time and attention on this website.

To understand the different user navigational patterns of each cluster, the averaged proportion of pages of each cluster in the online roadshow weblog and e-commerce weblog are plotted in Fig. 9 and Fig. 10, respectively. Both figures provide an initial understanding of the session cluster behavior before the user engagement assessment.

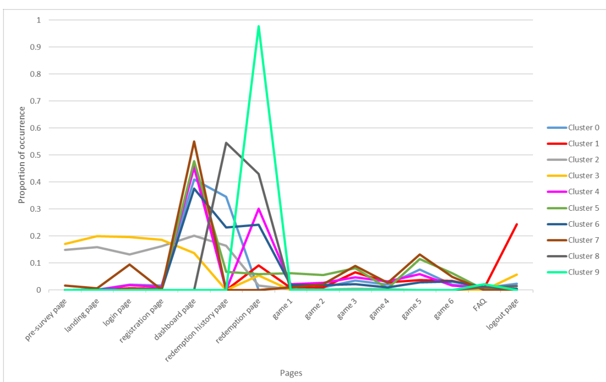

FIGURE 9. **Cluster profiles of online roadshow weblog.**

Fig. 9 leads to the following observation: Cluster 0 (blue) contains sessions that have high frequency on the redemption history page, which suggests that users in this session cluster only enter the online roadshow website to check their redemption history. Cluster 1 (red) contains users who

TABLE 12. **Interaction pattern of the session clusters in the online roadshow weblog.**

Cluster	Interaction pattern
Cluster 0	Existing users who visit the website to check their redemption history.
Cluster 1	Existing users who visit the website to redeem vouchers but log out shortly.
Cluster 2	Newly registered users who visit the redemption history page after creating an account.
Cluster 3	Newly registered users who visit the redemption page after creating an account.
Cluster 4	Existing users who visit the website to redeem vouchers, and occasionally play games.
Cluster 5	Existing users who visit the website to play games and redeem the vouchers after playing them.
Cluster 6	Existing users who visit the website to redeem vouchers, check their redemption history, and occasionally play games.
Cluster 7	Existing users who visit the website solely to play games.
Cluster 8	Existing users who visit the website to redeem vouchers and check their redemption history.
Cluster 9	Existing users who visit the website solely to redeem voucher.

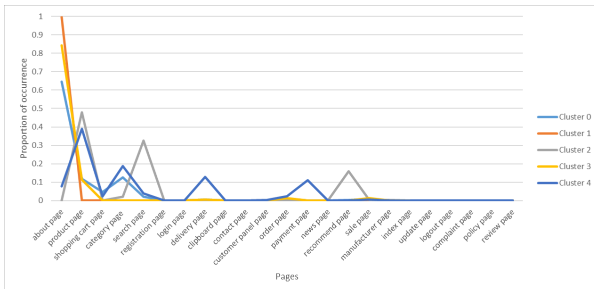

FIGURE 10. **Cluster profiles of e-commerce weblog.**

leave the website shortly after entering, which is reflected in the high frequency of the logout page. In Cluster 2 (grey) and Cluster 3 (yellow), the frequent pages are pre-survey page, landing page, login page, and registration page. This shows that both clusters contain newly registered users of the website. The main difference between both clusters is that Cluster 2 has higher frequency on the redemption history page, meanwhile Cluster 3 has higher frequency on the redemption page. Cluster 4 (magenta) has a high frequency on the redemption page, showing that the users' main aim in visiting the website is to redeem vouchers. In Cluster 5 (green) and Cluster 7 (brown), the visit frequency of the 6 game pages is high, which suggests a game playing behavior. However, there is a higher frequency on the redemption page and redemption history page in Cluster 5, depicting that users of this cluster have played the games and continued to redeem vouchers in a single visit. In Cluster 6 (dark blue), the users mainly visited the website to redeem vouchers and checked their redemption history. Cluster 8 (black) is similar to Cluster 6, but the users did not visit other web pages besides the redemption page and redemption history page.

TABLE 13. Interaction pattern of the session clusters in the e-commerce weblog.

Cluster	Interaction pattern
Cluster 0	Users who visit the website want to know more information about the website, and search for products based on its category.
Cluster 1	Users who visit the website to solely to know more information about the website.
Cluster 2	Users who visit the website to search for products with the search function and recommend products to others.
Cluster 3	Users who visit the website to know more information about the website and search for products.
Cluster 4	Users who visit the website to search for products based on its category, make payment, and check the delivery status of the purchased product.

In Cluster 9 (spring green), the users solely visited the website to redeem vouchers. Table 12 summarizes the navigation behavior of each session cluster on the online roadshow website.

Fig. 10 leads to the following observation: Cluster 0 (blue) contains sessions that have high frequency on the about page, category page, and product page. This suggests that the users in this session cluster entered the website to find more information about it and browsed through different categories of products listed on the website. In Cluster 1 (orange), users entered the website solely to discover more information about it. The most frequently visited pages in Cluster 2 (grey) are product page, search page, and recommend page. This shows that the users in this cluster have focused to search products and to recommend products to others. In Cluster 3 (yellow), users discovered more information about the website and browsed products to purchase. Cluster 4 (dark blue) has a high page frequency on the product page, category page, delivery page, and payment page. This shows that the users visited the website to browse through product categories, make payment for the products, and check the product delivery status after purchasing. Table 13 summarizes the navigation behavior of each session cluster on the e-commerce website.

Table 14 and 15 show the engagement levels of each UE metrics and attention scores of the session clusters in the online roadshow weblog and e-commerce weblog, respectively. The clusters in both tables are sorted in a descending order based on their attention score. Within the tables, the 'pages' column presents the top three frequently visited pages along with their averaged proportions in each cluster. The pages listed in the stated column follow the 'page name' in Table 3 and 4.

Based on Table 14, the session cluster with the highest attention score is Cluster 0, meanwhile, the session cluster with the lowest attention score is Cluster 1. Cluster 7 has the smallest size, with only 132 sessions (5.6%=132/2,361). This shows that only a few sessions contain users who visited the website solely to play games. Cluster 0 is the largest cluster, with 20.5% (485/2,361) of the sessions being assigned to it. This shows that the website received high attention from many users who return to the website to check their redemption status and redemption history. Cluster 4, Cluster 5, and Cluster 7 are session clusters that have game page as one of their top frequently visited pages. Among the three clusters listed, Cluster 5 has the highest attention score, meanwhile Cluster 4 and Cluster 7 have the medium attention level. Based on the UE metrics of Cluster 5, it received very low daily traffic despite having a high hourly activeness and high hourly traffic. This shows that there is only a very low number of visits from the users who redeem vouchers after playing games daily. However, this does not mean that the website received low attention from the users, as the website achieved a high engagement level for the three other UE metrics except daily traffic. Cluster 4, Cluster 6, Cluster 8 and Cluster 9 have medium attention level (2.25 to 2.50). The similarity between these four clusters is that the most frequently visited page is the redemption page. This indicates that the online roadshow website received medium attention from users when they were redeeming vouchers. Cluster 4 and Cluster 6 have the same attention score of 2.50. Both clusters have a high number of visits on the redemption page and a low number of visits on the game pages. However, the engagement levels of UE metrics show that both clusters are engaged in different aspects. Based on the metrics, users in Cluster 6 are very active on the website at any hour, but they are not active on the website daily. Meanwhile, users in Cluster 4 are moderately active on the website hourly and daily.

As Cluster 9 has equal amount of *low engagement* and *medium engagement* in the daily activeness metric, this means that there is both a low and moderate number of clicks collected daily from this cluster. The attention score of Cluster 9 is higher than Cluster 8 because of its high hourly traffic metric. The website received a high number of visits hourly in Cluster 9, while the website received a low number of visits hourly in Cluster 8. Compared to the other clusters in the medium attention category, Cluster 3 and Cluster 7 have the lowest attention score. Based on the hourly activeness and daily activeness metrics, there is a low number of users who are active on the website solely to play games at both hourly and daily in Cluster 7. On the other hand, the engagement level of the daily activeness metric in Cluster 3 is much lower than in Cluster 7. The difference in engagement level shows that the attention received by the website comes from different aspects (hourly activeness and daily activeness). The main difference between Cluster 3 and Cluster 7 are: Cluster 3 has a higher hourly user interaction than Cluster 7, however, Cluster 3 has a lower daily user interaction than Cluster 7. The online roadshow website received less attention in Cluster 1 and Cluster 2. The attention score of Cluster 1 is lower than Cluster 2, meaning that the website received the least attention and involvement from users who log out shortly during their visit. Compared to Cluster 1, Cluster 2 has a higher attention score as the users are newly registered and visited the redemption history page after creating an account on the website. The high page visit of the redemption history

TABLE 14. Cluster characteristics of online roadshow website.

Cluster	Page	Engagement level of UE metrics				Attention score
		Hourly activeness	Hourly traffic	Daily activeness	Daily traffic	
Cluster 0 (N=485)	dashboard page (0.410) redemption history page (0.346) game 5 (0.076)	High	High	High	Medium	3.75 (High)
Cluster 5 (N=232)	dashboard page (0.478) game 5 (0.115) game 3 (0.081)	High	High	High	Very low	3.25
Cluster 6 (N=264)	dashboard page (0.376) redemption page (0.242) redemption history page (0.232)	High	Medium	Very low	Low	2.50 (Medium)
Cluster 4 (N=257)	dashboard page (0.453) redemption page (0.302) game 5 (0.059)	Medium	High	Medium	Low	2.50
Cluster 9 (N=289)	redemption page (0.978) FAQ page (0.022)	Low	High	Low/Medium	Low	2.38
Cluster 8 (N=216)	redemption history page (0.546) redemption page (0.431) logout page (0.014)	Low	Low	Medium	Low	2.25
Cluster 3 (N=144)	landing page (0.200) login page (0.197) registration page (0.187)	Medium	Low	Very low	Low	2.00
Cluster 7 (N=132)	dashboard page (0.550) game 5 (0.131) login page (0.094)	Low	Low	Low	Low	2.00
Cluster 2 (N=150)	dashboard page (0.202) redemption history page (0.164) registration page (0.162)	High/Medium/Low/ Very low	Low	Very low	Low	1.38 (Low)
Cluster 1 (N=192)	dashboard page (0.464) logout page (0.243) redemption page (0.091)	Very low	Very low	Very low	Very low	1.00

page indicates that a few newly registered users participate in the online roadshow as they are attracted to the redemption gift. Although Cluster 2 has a similar pattern to Cluster 3, Cluster 2 has a low attention score as the users have low interaction with the website at any hour.

In Table 15, the session cluster with the highest attention score is Cluster 1, which is the largest cluster as well. Cluster 2 has the smallest cluster size and the lowest attention score among the five clusters. The engagement levels of the UE metrics of Cluster 1 are high, showing that e-commerce website received both high interaction and user visits in this cluster at any hour and day. Therefore, Cluster 1 has the topmost attention score of 4.00. Besides Cluster 1, Cluster 3 and Cluster 0 have a high attention score. The similarity among these three clusters is that the most frequently visited page is the about page. This means that the e-commerce website receives high attention from the users when they are visiting the about page to know more information about the website. As the averaged proportion of the about page decreases, the attention score of the cluster decreases as well. The attention score of Cluster 3 is lower than Cluster 1 since the engagement levels of its hourly activeness, daily activeness and daily traffic are medium. Cluster 0 has the lowest attention score among the three high attention clusters. Although the engagement levels of its hourly activeness and daily activeness are high, the daily traffic metric is low. This shows that the users have high interaction when navigating the website despite visiting the website infrequently daily. Cluster 4 and Cluster 2 have medium and low attention scores, respectively. The most frequently visited page of both clusters is the product page, which is the page that the users are expected to visit. This shows that the website received less attention from users who enter the website to search for products.

TABLE 15. Cluster characteristics of e-commerce website.

Cluster	Page	Engagement level of UE metrics				Attention score
		Hourly activeness	Hourly traffic	Daily activeness	Daily traffic	
Cluster 1 (N=42,691)	about page (0.995) order page (0.002) category page (0.0002)	High	High	High	High	4.00 (High)
Cluster 3 (N=20,448)	about page (0.842) product page (0.114) order page (0.014)	Medium	High	Medium	Medium	3.25
Cluster 0 (N=17,263)	about page (0.644) category page (0.126) product page (0.118)	High	Medium	High	Low	3.00
Cluster 4 (N=17,395)	product page (0.389) category page (0.187) delivery page (0.128)	Low	Medium	Low	Low	2.25 (Medium)
Cluster 2 (N=7,078)	product page (0.479) search page (0.325) recommend page (0.160)	Low	Very low/Low	Very low	Very low	1.38 (Low)

VI. DISCUSSION

The Cluster-N-Engage framework offers a robust alternative to the conventional methods of measuring user engagement. Conventionally, using questionnaires and interviews is a more straightforward method to measure user engagement of a website. However, this method is very subjective as it requires users to provide their opinions towards the website [13]. In addition, users may have different interpretations of the questions asked, or their feelings may affect the answers to the questions, which can further affect the reliability of the data collected. Another method to evaluate user engagement is to use physiological measures, such as camera and eye trackers. Although this approach solves the subjectivity issue, it is more invasive, as it requires a physical device to be placed near the users to record their movement.

In the Cluster-N-Engage framework, website engagement is evaluated by extracting the user navigational behavior from its weblog via web usage mining [40], [41]. Compared to questionnaires and physiological data, web usage mining can capture user behaviors automatically with no direct interaction with the users. This is beneficial as users can navigate the website naturally, with no awareness of being observed.

The Cluster-N-Engage framework presents how the four proposed user engagement metrics, which are hourly activeness, hourly traffic, daily activeness, and daily traffic, can effectively describe the user navigational behavior of the session clusters. Each metric illustrates the engagement level of a session cluster on the website from different aspects. The introduction of the proposed metrics provides a clearer understanding of the user navigational behavior of each cluster. For example, if a session cluster exhibits high engagement across all four metrics, it suggests that the users within that cluster are actively interacting with the website, navigating it frequently on an hourly and daily basis. These metrics reflect user attention towards a website. Lehmann et al. [8] proposed a set of engagement metrics that reflect the popularity, user activity, and user loyalty of a website. The findings of their research conclude that the engagement level of the metrics is highly influenced by the type of websites. For instance, the popularity metric might indicate a low engagement value for an interest-based website [8], [42] because of a relatively small number of users. However, this should not be interpreted as a low engagement, as interest-based websites are designed for a small group of users of a specific interest. In the Cluster-N-Engage framework, the engagement level of each metric is not affected by the website types. The application of a quartile formula in each metric individually makes sure that the categorization of engagement levels is unique to each specific website. The flexibility of the proposed frameworks allows it to be applied across various types of websites for measuring user engagement.

Many existing frameworks attempted to use a single score to measure the overall engagement of the website. This engagement score is calculated by combining the results of multiple user engagement metrics. The biggest difference between the proposed framework and the existing framework that is the combination of clustering and attention score (which is the engagement score). The user navigational behavior of the weblog is first clustered before being evaluated using the four proposed user engagement metrics. Based on the metrics, the attention score of each session cluster within the website is computed. The attention score acts as a single score that shows the user attention level of each session cluster, leading to the identifying of navigational behavior that contribute to a good engagement in a website. By examining the attention score of the session clusters, organizations can investigate the factors that influence user attention on the website. Different clusters of navigational behavior have distinct levels of user attention. Therefore, organizations can identify the navigational behavior that exhibit high user attention on the website, allowing them to understand the factors driving user attention. Capturing user attention on a website is crucial for creating a good impression among the users, leading to effective advertising. The advertising effectiveness of the website highly depends on the objective that the organizations have on the website. A high attention score in session clusters that involves the target actions of the website means the website is effective in achieving the website aim. With the online roadshow, the main objective is to attract users to play games to redeem vouchers. The products-specific advertising elements are embedded into the games to engage the users in the online roadshow website [26]. Organizations would expect that the session clusters involving games to have high attention score to conclude the website as successful. Clusters that have a high number of visits of game pages are Cluster 4 (medium attention), Cluster 5 (high attention), and Cluster 7 (medium attention). This means that the online roadshow website is effective in achieving its objective as it received high and medium attention from clusters that involve the target actions. The e- commerce website aims to attract users to find products and purchase them. The clusters that are related to these target actions are Cluster 2 (medium attention) and Cluster 4 (low attention) because of the high number of page visits on the product page. This shows that the e-commerce website is least significant in achieving its objective as it received less attention from these clusters. Therefore, developers can leverage this knowledge to make some enhancements on the e-commerce website.

While this study offers valuable insights into user engagement on websites, it is important to acknowledge certain limitations. First, the Cluster-N-Engage framework's effectiveness heavily relies on the completeness of the weblog collected. Inaccuracies or gaps in weblogs, such as missing clicks and incomplete session record, could impact the calculation of the attention score. The study examines two website types, namely the online roadshow website and the e-commerce website, to validate the applicability of the framework. In the future, more engagement metrics can be added to framework to capture other dimensions of user attention. Nonetheless, the Cluster-N-Engage framework offers valuable insights into website engagement and opens the opportunity for further research to address these constraints. products.

VII. CONCLUSION

The Cluster-N-Engage framework presented in this paper aims to take a deeper dive into user engagement on websites. User navigational behavior is first extracted from the weblog by performing web usage mining. Web session clustering is then performed on the online roadshow and e-commerce weblog to group session clusters according to their similarity. After the clustering process, four user engagement metrics, namely hourly activeness, hourly traffic, daily activeness, and daily traffic are used to measure the engagement level of each session cluster. To make decision-making easier among organizations, a single score is aggregated from multiple user engagement metrics to provide the overall engagement of the website. According to the engagement levels of each user engagement metric, the attention score is formulated to show the amount of attention that the website received from the users of each session cluster. Besides that, the attention score can provide insights into the success of the website in achieving its objective.

The future work in this study involves analyzing the transition of session clusters in users. Web sessions of a user can belong to many session clusters. By analyzing the cluster transitions in users, more user navigational patterns can be extracted to understand the level of user attention varies across different clusters.

REFERENCES

[1] P. Kavassalis, S. Lelis, M. Rafea, and S. Haridi, "What makes a web site popular?" *Commun. ACM*, vol. 47, no. 2, pp. 50–55, Feb. 2004, doi: 10.1145/966389.966415.

[2] L. C. Schaupp, W. Fan, and F. Belanger, "Determining success for different website goals," in *Proc. 39th Annu. Hawaii Int. Conf. Syst. Sci. (HICSS)*, Kauia, HI, USA, 2006, pp. 1–10, doi: 10.1109/HICSS.2006.122.

[3] J. Jankowski, P. Kazienko, J. Wątróbski, A. Lewandowska, P. Ziemba, and M. Zioło, "Fuzzy multi-objective modeling of effectiveness and user experience in online advertising," *Expert Syst. Appl.*, vol. 65, pp. 315–331, Dec. 2016, doi: 10.1016/j.eswa.2016.08.049.

[4] S.-J. Lee, W.-N. Lee, H. Kim, and P. A. Stout, "A comparison of objective characteristics and user perception of web sites," *J. Interact. Advertising*, vol. 4, no. 2, pp. 61–75, Mar. 2004, doi: 10.1080/15252019.2004.10722088.

[5] P. J. Q. Ferreira, P. Rita, D. Morais, P. J. Rosa, J. Oliveira, P. Gamito, N. Santos, F. Soares, and C. Sottomayor, "Grabbing attention while reading website pages: The influence of verbal emotional cues in advertising," *J. Eye Tracking, Vis. Cogn. Emotion*, vol. 1, no. 1, pp. 64–68, 2011. [Online]. Available: https://recil.ensinolusofona.pt/handle/10437/2312

[6] M. U. Özmen, "Online information retrieval behaviour and economics of attention," *Online Inf. Rev.*, vol. 39, no. 6, pp. 779–794, Oct. 2015, doi: 10.1108/OIR-05-2015-0139.

[7] A. Mollen and H. Wilson, "Engagement, telepresence and interactivity in online consumer experience: Reconciling scholastic and managerial perspectives," *J. Bus. Res.*, vol. 63, nos. 9–10, pp. 919–925, Sep. 2010, doi: 10.1016/j.jbusres.2009.05.014.

[8] J. Lehmann, M. Lalmas, E. Yom-Tov, and G. Dupret, "Models of user engagement," in *Proc. 20th Int. Conf. User Modeling Adaptat. Pers. (UMAP)*, Montreal, QC, Canada, 2012, pp. 164–175, doi: 10.1007/978-3-642-31454-4_14.

[9] S. Attfield, G. Kazai, and M. Lalmas, "Towards a science of user engagement (position paper)," in *Proc. WSDM Work. User Modeling Web Appl.*, Hong Kong, Feb. 2011, pp. 1–8. [Online]. Available: http://www.dcs.gla.ac.uk/~mounia/Papers/engagement.pdf

[10] H. L. O'Brien, P. Cairns, and M. Hall, "A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form," *Int. J. Hum.-Comput. Stud.*, vol. 112, pp. 28–39, Apr. 2018, doi: 10.1016/j.ijhcs.2018.01.004.

[11] F. D. Saiful Bahry, M. Masrom, and M. N. Masrek, "Measuring validity and reliability of website credibility factors in influencing user engagement questionnaire," *Int. J. Web Inf. Syst.*, vol. 17, no. 1, pp. 18–28, Jan. 2021, doi: 10.1108/IJWIS-08-2020-0050.

[12] R. Craig Lefebvre, Y. Tada, S. W. Hilfiker, and C. Baur, "The assessment of user engagement with eHealth content: The eHealth engagement scale," *J. Comput.-Mediated Commun.*, vol. 15, no. 4, pp. 666–681, Jul. 2010, doi: 10.1111/j.1083-6101.2009.01514.x.

[13] M. Slater, "Measuring presence: A response to the witmer and singer presence questionnaire," *Presence Teleoperators Virtual Environ.*, vol. 8, no. 5, pp. 560–565, Oct. 1999, doi: 10.1162/105474699566477.

[14] K. Sharma, G. Shrivastava, and V. Kumar, "Web mining: Today and tomorrow," in *Proc. 3rd Int. Conf. Electron. Comput. Technol.*, Kanyakumari, India, Apr. 2011, pp. 399–403, doi: 10.1109/ICECTECH.2011.5941631.

[15] P. M. Bharti and T. J. Raval, "Improving web page access prediction using web usage mining and web content mining," in *Proc. 3rd Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Coimbatore, India, Jun. 2019, pp. 1268–1273, doi: 10.1109/ICECA.2019.8821950.

[16] R. Cooley, B. Mobasher, and J. Srivastava, "Web mining: Information and pattern discovery on the world wide web," in *Proc. 9th IEEE Int. Conf. Tools Artif. Intell.*, Newport Beach, CA, USA, Nov. 1997, pp. 558–567, doi: 10.1109/TAI.1997.632303.

[17] O. Etzioni, "The world-wide web: Quagmire or gold mine?" *Commun. ACM*, vol. 39, no. 11, pp. 65–68, Nov. 1996, doi: 10.1145/240455.240473.

[18] Z. Y. Lim, L. Y. Ong, and M. C. Leow, "A review on clustering techniques: Creating better user experience for online roadshow," *Future Internet*, vol. 13, no. 9, pp. 1–21, 2021, doi: 10.3390/fi13090233.

[19] P. Parihar, J. Dawra, and V. Sahay, "The role of customer engagement in the involvement-loyalty link," *Marketing Intell. Planning*, vol. 37, no. 1, pp. 66–79, Feb. 2019, doi: 10.1108/MIP-11-2017-0318.

[20] L. Liang, "Essays on user acquisition, retention, experience, engagement and purchase decisions of a hybrid mobile wallet," Ph.D. dissertation, Fac. Graduate School, Univ. Texas, Austin, TX, USA, 2017. [Online]. Available: http://hdl.handle.net/2152/63040

[21] D. P. Sakas and D. P. Reklitis, "The impact of organic traffic of crowdsourcing platforms on airlines' website traffic and user engagement," *Sustainability*, vol. 13, no. 16, pp. 1–22, 2021, doi: 10.3390/su13168850.

[22] F. D. S. Bahry, M. Masrom, and M. N. Masrek, "Website evaluation measures, website user engagement and website credibility for municipal website," *ARPN J. Eng. Appl. Sci.*, vol. 10, no. 23, pp. 18228–18238, Dec. 2015. [Online]. Available: https://www.researchgate.net/profile/Farrah-Diana-Saiful-Bahry/publication/290460387_Website_evaluation_measures_website_user_engagement_and_website_credibility_for_Municipal_website/links/56a981bc08ae2df8216534a4/Website-evaluation-measures-website-user-engagement-and-website-credibility-for-Municipal-website.pdf

[23] E. T. Peterson and J. Carrabis, "Measuring the immeasurable: Visitor engagement," Web Anal. Demystified, Camas, WA, White Paper 14, 2008, vol. 16. [Online]. Available: https://www.verticalstudio.com/hs-fs/hub/74398/file-15425996-pdf/docs/web_analytics_demystified_and_nextstage_global_-_measuring_the_immeasurable_-_visitor_engagement.pdf

[24] D. Lagun and M. Lalmas, "Understanding user attention and engagement in online news reading," in *Proc. 9th ACM Int. Conf. Web Search Data Mining*, San Francisco, CA, USA, Feb. 2016, pp. 113–122, doi: 10.1145/2835776.2835833.

[25] S. Fedushko, T. Ustyianovych, Y. Syerov, and T. Peracek, "User-engagement score and SLIs/SLOs/SLAs measurements correlation of e-business projects through big data analysis," *Appl. Sci.*, vol. 10, no. 24, pp. 1–16, 2020, doi: 10.3390/app10249112.

[26] K. R. Leow, M. C. Leow, and L. Y. Ong, "Online roadshow: A new model for the next-generation digital marketing," in *Proc. Future Technol. Conf. (FTC)*, Vancouver, BC, Canada, Oct. 2021, pp. 994–1005, doi: 10.1007/978-3-030-89906-6_64.

[27] Z. Y. Lim, L. Y. Ong, and M. C. Leow, "Experimental study on predictive modeling in the gamification marketing application," in *Proc. 34th Int. Conf. Ind. Eng. Appl. Appl. Intell. Syst. (IEA/AIE)*, Kuala Lumpur, Malaysia, Jul. 2021, pp. 379–390, doi: 10.1007/978-3-030-79463-7_32.

[28] G. Chodak, G. Suchacka, and Y. Chawla, "EClog: HTTP-level e-commerce data based on server access logs for an online store," V1, Distrib. Harvard Dataverse, 2020 doi: 10.7910/DVN/Z834IK.

[29] G. Chodak, G. Suchacka, and Y. Chawla, "HTTP-level e-commerce data based on server access logs for an online store," *Comput. Netw.*, vol. 183, pp. 1–5, Dec. 2020, doi: 10.1016/j.comnet.2020.107589.

[30] V. Sujatha and Punithavalli, "Improved user navigation pattern prediction technique from web log data," *Proc. Eng.*, vol. 30, pp. 92–99, Jan. 2012, doi: 10.1016/j.proeng.2012.01.838.

[31] M. Srivastava, R. Garg, and P. K. Mishra, "Analysis of data extraction and data cleaning in web usage mining," in *Proc. Int. Conf. Adv. Res. Comput. Sci. Eng. Technol. (ICARCSET)*, Unnao, India, Mar. 2015, pp. 1–6, doi: 10.1145/2743065.2743078.

[32] J. X. Yu, Y. Ou, C. Zhang, and S. Zhang, "Identifying interesting customers through web log classification," *IEEE Intell. Syst.*, vol. 20, no. 3, pp. 55–59, May 2005, doi: 10.1109/MIS.2005.47.

[33] K. Selvakumar, L. S. Ramesh, and A. Kannan, "Enhanced K-means clustering algorithm for evolving user groups," *Indian J. Sci. Technol.*, vol. 8, no. 24, pp. 1–8, Sep. 2015, doi: 10.17485/ijst/2015/v8i24/80192.

[34] S. Asadianfam and M. Mohammadi, "Identify navigational patterns of web users," *Int. J. Comput. Technol.*, vol. 1, no. 1, pp. 1–8, 2014.

[35] A. Guerbas, O. Addam, O. Zaarour, M. Nagi, A. Elhajj, M. Ridley, and R. Alhajj, "Effective web log mining and online navigational pattern prediction," *Know.-Based Syst.*, vol. 49, pp. 50–62, Sep. 2013, doi: 10.1016/j.knosys.2013.04.014.

[36] H. Humaira and R. Rasyidah, "Determining the appropiate cluster number using elbow method for K-means algorithm," in *Proc. 2nd Workshop Multidisciplinary Appl. (WMA)*, Padang, Indonesia, 2020, pp. 1–8, doi: 10.4108/eai.24-1-2018.2292388.

[37] A. S. Shirkhorshidi, S. Aghabozorgi, and T. Y. Wah, "A comparison study on similarity and dissimilarity measures in clustering continuous data," *PLoS One*, vol. 10, no. 12, pp. 1–20, 2015, doi: 10.1371/journal.pone.0144059.

[38] Z. Ansari, M. F. Azeem, W. Ahmed, and A. V. Babu, "Quantitative evaluation of performance and validity indices for clustering the web navigational sessions," *World Comput. Sci. Inf. Technol. J.*, vol. 1, no. 5, pp. 217–226, 2011.

[39] Z.-Y. Lim, L.-Y. Ong, M.-C. Leow, T.-W. Lee, and Q.-M. Tay, "Understanding user behaviour with web session clustering and user engagement metrics," in *Proc. 19th IEEE Int. Colloq. Signal Process. Appl. (CSPA)*, Kedah, Malaysia, Mar. 2023, pp. 19–24, doi: 10.1109/CSPA57446.2023.10087488.

[40] G. Neelima and S. Rodda, "Predicting user behavior through sessions using the web log mining," in *Proc. Int. Conf. Adv. Human Mach. Interact. (HMI)*, Kodigehalli, India, Mar. 2016, pp. 1–5, doi: 10.1109/HMI.2016.7449167.

[41] G. Castellano, A. M. Fanelli, and M. A. Torsello, "Log data preparation for mining web usage patterns," in *Proc. Int. Conf. Appl. Comput. (IADIS)*, 2007, pp. 371–378.

[42] D. Li, Q. Lv, X. Xie, L. Shang, H. Xia, T. Lu, and N. Gu, "Interest-based real-time content recommendation in online social communities," *Knowl.-Based Syst.*, vol. 28, pp. 1–12, Apr. 2012, doi: 10.1016/j.knosys.2011.09.019.

ZHOU-YI LIM received the bachelor's degree in computer science (artificial intelligence) from Multimedia University, Malaysia, where she is currently pursuing the master's degree. Her current research interests include behavior and web usage mining.

LEE-YENG ONG (Senior Member, IEEE) received the Ph.D. degree in computer vision from Multimedia University, Malaysia, in 2020. She is currently a Senior Lecturer with the Faculty of Information Science and Technology, Multimedia University. Her current research interests include object tracking, data science, and big data analytics.

MENG-CHEW LEOW (Senior Member, IEEE) received the D.Phil. degree from Multimedia University. His current research interests include game-based learning, system science, practical spirituality, and philosophy, specifically in role-playing game-based learning.