ASEAN Engineering Journal

Full Paper

SOCIO-ECONOMIC DEVELOPMENT INDICES AND THEIR REFLECTION ON INTERNET PERFORMANCE IN ASEAN COUNTRIES

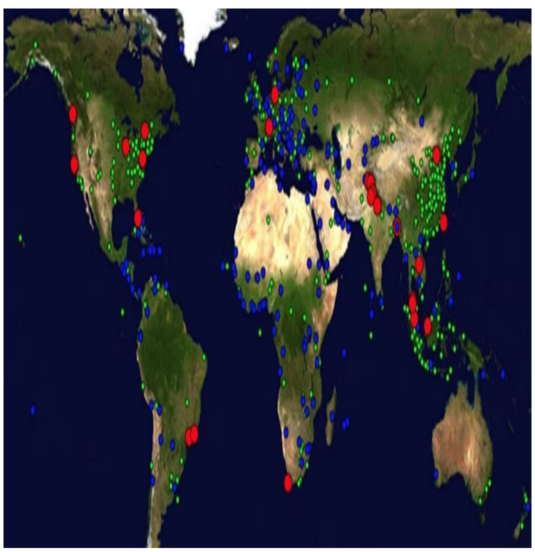
Adib Habbal[a], Emmanuel O.C. Mkpojiogu[b,d*], Les Cottrell[c], Bebo White[c], Suhaidi Hassan[d]. Faisal Zulhumadi[d]

[a]Computer Engineering Department, Faculty of Engineering, Karabuk University, 78050, Karabuk, Türkiye
[b]Department of Software Engineering, Veritas University, Abuja, Nigeria
[c]SLAC National Accelerator Laboratory, Stanford University, Menlo Park, CA 94025, USA,
[d]School of Computing, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia

Article history
Received
25 May 2022
Received in revised form
21 February 2023
Accepted
29 August 2023
Published online
29 February 2024

*Corresponding author
mkpojiogue@veritas.edu.ng

Graphical abstract



Abstract

Despite Internet use rapidly accelerating in ASEAN countries, its penetration rate across member countries varies from 84.45% of the population in Singapore to roughly 21.87% in Laos. This digital divide portends profound consequences on the social-economic development of the region. Therefore, this article describes an Internet performance study conducted within ASEAN countries using actual Internet performance data collected from 2000 to 2019 generated by the PingER Project. The results showed that the pattern of Internet performance (IP) is that the most developed countries have the best Internet performance, whereas the least developed ones have the lowest Internet performance. These Internet performance data were then compared and analyzed against several selected social-economic development indices in order to observe any trends and establish observable relationships. Initial inspection indicated a possible relationship between Internet performance and the indices, which was then statistically tested further by correlation and regression. Next, the relationship was then represented by a regression model, which was then validated through R2 and graphical residual analysis. As a result, this study has proposed a novel model that provides an insight into the influence of social-economic development indices on Internet performance in ASEAN countries.

Keywords: Development Index, Internet Performance, Internet Socio-economic Model, PingER, ASEAN

1.0 INTRODUCTION

Association of Southeast Asian Nations (ASEAN), established in August 1967, is a geo-political and economic organization made up of 10 South East Asian member countries namely Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, and Vietnam. One of the main objectives of this organization is to promote inter-governmental cooperation and accelerate the economic growth amongst its members. The association is meant to play a key role in the prosperity, peace, and geographical stability of member nations [1]. ASEAN is one of the fastest growing sub-regions in the world with a projected growth rate of 5.4% per annum from 2014 to 2018 [2]. In addition, it is projected that this sub-region, as a single entity, would be the seventh largest global economy with a combined Gross Domestic Product (GDP) of $2.4 trillion in 2013 [2].

Researchers recognize ASEAN countries as a sub-region in the world with the most potential for growth and development [3].

ASEAN countries differ in their income distribution as well as Internet penetration [4]. Higher Internet penetration is associated with a higher GDP. There is a strong association between ICT diffusion and GDP growth [4]. Studies reveal that Singapore has the highest Internet penetration rate in the sub-region. While Srinuan et al. [4] report that Malaysia (55.70%), Brunei (48.82%), Thailand (21%), and Vietnam (20.45%) have a higher Internet usage rate than other ASEAN countries. Cambodia, Myanmar, and Laos have the lowest Internet penetration rate among the ASEAN countries. The estimated Internet penetration in 2017 is as follows: Brunei moves a step ahead of Singapore and Malaysia with the highest Internet usage rate (90%); while Singapore and Malaysia are placed in the second and third positions with (84.45%) and (80.14%) respectively; while Philippines (55.5%), Thailand (52.89%), and Vietnam (46.5%) have managed to increase the Internet penetration rate significantly in 2017 as compared to Indonesia (32.29%) and Cambodia (34%). In contrast, Myanmar (25.7%) and Laos (21.87) maintain their position with the lowest Internet penetration rate among the ASEAN countries [5].

In terms of a competitive economy, only five ASEAN countries (Malaysia, Indonesia, Singapore, the Philippines, and Vietnam [4]) have a fully competitive economy. The remaining countries have a partly competitive and partly monopolistic economic structure. However, only Singapore, Malaysia, Brunei, the Philippines, and Thailand have independent regulators [4]. To be and remain competitive, ASEAN member countries believe that they must foster deeper regional integration [1]. However, ASEAN economies are characterized by vast differences, for instance, Singapore has a GDP per capita that is >60 times higher than Myanmar's. Thus, narrowing development gaps and growth that is equitable and inclusive are very critical and crucial for the successful integration of ASEAN countries [1]. One way of addressing this challenge is enhancing connectivity, part of which is Internet connectivity [1]. Broadband Internet access is considered by many economists and policymakers as the way to achieve a knowledge-based economy [6].

To deepen ASEAN regional integration, Internet connectivity is recognized as an important tool [1]. However, currently, Internet penetration across these countries varies from 90% of the population in Brunei to little more than 21.87% in Laos. This digital divide portends profound consequences [1]. Therefore, it is clearly necessary to unravel the Internet performance situation across ASEAN countries. On the other hand, it is a challenging task to classify countries according to their development. One has to carefully select suitable metrics related to the development and then measure them. There are then costs and practicality concerns for instance; which metrics can be measured, how useful they are, how well defined they are, how they vary over time, and the actual measurement process. Gathering these metrics takes time, so they are often made available at widely separated intervals.

In this research, the Internet performance of ASEAN countries and how it associates with several main socio-economic development indices were investigated. Specifically, the contributions of this research include:

- The collection of real data measurements about Internet performance of ASEAN countries from 2000-2019 using the Ping End-to-end Reporting (PingER) methodology [7] (Section III).
- The evaluation of Internet performance of ASEAN countries to reveal their current state. It is shown that ASEAN countries can be categorized into three groups according to their Internet performance (Section IV).
- The proposal of a novel model to understand the impact of socio-economic development indices on ASEAN countries' Internet performance (Section V).

The proposed model provides an insight into the influence of Gross Domestic Product (GDP), ICT Development Index (IDI), Human Development Index (HDI), fertility rate, and Happy Planet Index (HPI) indices on Internet performance. It is proven, through the developed regression model, that PingER end-to-end Internet measurements can be used as a good indicator of a country's development progress towards achieving a digital economy.

2.0 BACKGROUND

Network interconnections are diverse among ASEAN countries. In more developed countries like Singapore and Malaysia, network interconnection is more advanced than in the lower-income countries that have fewer IP-enabled networks. Also, Internet connectivity and performance status are not uniform among the 10 ASEAN countries. According to ISOC [1], for most ASEAN countries, Internet penetration values correlate with income level (based on the World Bank classification). However, this is untrue for Thailand and Indonesia. Their lower Internet penetration does not tally with their income level. This indicates that GDP per capita alone is insufficient in determining the Internet penetration level of a given ASEAN country. Certainly, a number of market and policy variables contribute to driving or constraining Internet adoption and hence, its penetration [1]. Table 1 shows the state of Internet connectivity in ASEAN countries. The very high wireless broadband penetration for Singapore for example could be because Singapore is a very rich country and hence, many users subscribe to multiple Subscriber Identity Module (SIM) cards and possibly multiple devices to take advantage of lower internet pricing available. This data is referenced from the precise and accurate source (Internet Society report 2015).

Table 1. Internet Connectivity Status for ASEAN Countries [1]

Cluster	Country	Income Group Type	Internet Users	Growth of Internet Users	Wireless Broadband Penetration
1	Singapore	High income	73%	1%	137%
	Malaysia	Upper middle	67%	3%	14%
	Brunei	High income	65%	6%	7%
2	Vietnam	Lower middle	44%	14%	22%
	Philippines	Lower middle	37%	43%	28%
	Thailand	Upper middle	29%	6%	53%
3	Indonesia	Lower middle	16%	22%	36%
	Laos	Low income	13%	46%	2%
	Cambodia	Low income	6%	59%	10%
	Myanmar	Low income	1%	38%	1%

One good indicator of a country's development towards an information economy era is the size of the Internet infrastructure [16]. However, calculating the number of users is a challenging task in developing countries since people use academic networks, share a single account, or access the Internet via telecentres, cyber cafés, and business services. Moreover, looking at the number of Internet users without taking into account the extent of use is inappropriate, since Internet usage varies from just sending/receiving a couple of emails a week, to spending hours a day streaming, browsing, or downloading. Therefore, new measures of Internet activity are needed to take these factors into consideration.

Another possible indicator to get an idea about a country's progress is to measure the international Internet bandwidth used by the country (i.e., the transmission capacity which is most often measured in Megabits per second, Mbps) [16]. In contrast to developed countries, 75-90% of Internet traffic in a developing country is international, so the size of its international traffic compared to population size provides a good indication of the Internet activity in a country.

Moreover, various organizations such as the World Bank, the International Telecommunication Union (ITU), and the United Nations Development Program (UNDP) have come up with several development indices such as the ICT Development Index (IDI) [10], Human Development Index (HDI) [11], and fertility rate [12], etc. However, to the best of our knowledge, none of the related works studied the direct relationship between Internet performance and those social-economic development indices. In other words, there is no existing model to show the relation between GDP, IDI, HDI, HPI, and FR on the one hand, and Internet performance on the other hand. This study is the first of its kind where a novel model is proposed using PingER end-to-end Internet measurements to examine the relationship between Internet performance and socio-economic development indices.

3.0 METHODOLOGY

This paper studies the Internet performance of ASEAN countries using PingER [7], which is the Internet End-to-end Performance Measurement (IEPM) project that monitors Internet-link end-to-end performance. Initially, the main goal behind the project was to monitor and understand the present Internet performance and to allocate resources to optimize performance between laboratories, universities, and institutes collaborating in the present and future experiments. The project was later expanded to include monitoring over 700 sites in more than 160 countries to characterize the digital divide. The monitored countries contain over 98% of the world's population as shown in Figure 1.

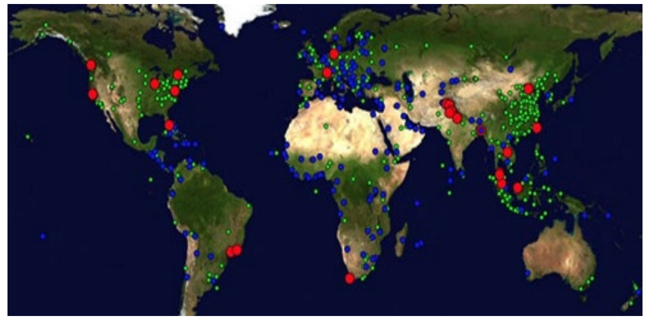


Figure 1. Locations of PingER MAs and remote sites as of December 2019. Red sites are MAs, Blue sites are beacons that are monitored by most MAs, and green sites are remote sites that are monitored by one or more MAs.

PingER has three host types as shown in Figure 2. The first is the monitoring host or Measurement Agent (MA). The MA is a computer on which the PingER monitoring software is deployed. MAs can be data center servers, desktops, laptops, or a single board computer with a minimum hardware requirement. The computers are installed with a Linux-based operating system. Furthermore, the computer must be connected to the Internet with a public IP address that is accessible from outside the given organization [7].

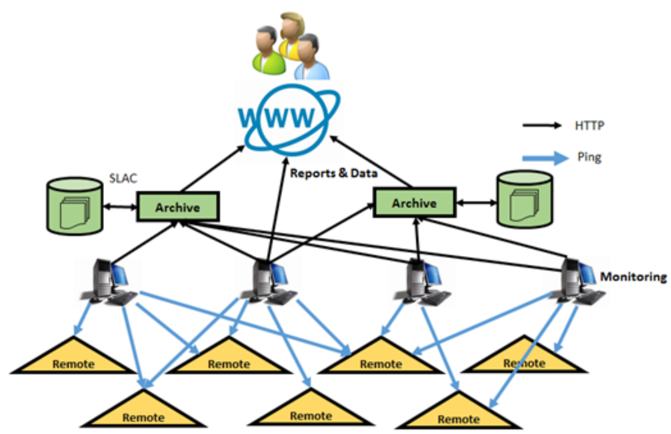


Figure 2. PingER Architecture

The second type is the remote host. Remote hosts are hosts monitored by MAs. They are usually servers with stable uptimes like web servers. For a remote host, no software or setup is required. The only thing required is that the host must respond to the ping. Currently, there are 27 monitored remote sites in Malaysia and 94 in Southeast Asia. The third type is the archive host. The archive host collects data from the MAs and also serves as a storage facility for the raw data collected. The primary archive host for the PingER project is at SLAC.

The PingER project is based on ping, a ubiquitous pre-installed program with the advantage of being lightweight when compared with other methods such as using SNMP or active probes [13]. In addition, PingER also has the advantage of measuring from an end user perspective instead of just monitoring the performance of the Internet backbone as done by Internet Service Providers (ISPs) [9]. So, the data collected through PingER reflects the end-user experience.

Measurements are carried out on the MA by sending to each remote host up to 30 pings at 1-second intervals until 10 responses are received. This is repeated every half hour. An XML file provides a set of remote hosts to ping. The first packet is usually discarded since it includes delays due to priming caches and other initialization procedures. Data is collected every day, with the use of HTTP, from the MAs to the archive host at SLAC. The RTTs for each set of pings are recorded. In addition, the archive host also carries out an analysis of the raw data captured and generates reports. Ping's ubiquity and ease of use make it suitable for widespread Internet monitoring, especially in less developed regions of the world.

Internet Performance Metrics: PingER data is objective and can be gathered instantly and thus analyzed more quickly and frequently. The PingER analyzed data consists of about 16 metrics. The most important ones are; packet loss, unreachability, Round-Trip Time (RTT), jitter, and throughput [14]. We used PingER analytical and measurement tools to obtain the Internet performance from 2000 to 2019 for each ASEAN country. Three Internet performance metrics, namely throughput, RTT, and packet loss are considered in this study to evaluate the Internet performance in ASEAN, as shown in Table 2. Furthermore, we selected PingER normalized throughput in 2015 to represent the Internet performance while studying the relationship with development indices. The PingER-derived throughput is calculated based on the formula (1) and it is measured from SLAC in North America so as to have the same point of reference for all countries.

Table 2. PingER metrics and their definitions [14]

Metric	Description
Packet Loss	A packet is considered lost if no ICMP echo reply is received at the originating node within 20 seconds from the time it sent the corresponding ICMP request packet.
Round-Trip Time (RTT)	This is the elapsed time between the sending of an ICMP echo request packet to a remote node and the receiving of a corresponding ICMP echo reply back to the origination host.
Throughput	Throughput is computed using the Mathis Formula [15]: $$Rate < \frac{MSS}{RTT} * \frac{1}{\sqrt{p}} \quad (1)$$ where MSS is the maximum segment size, RTT is the round-trip time, and p is the probability of packet loss.

Social Development Indices: Several development indices were proposed to look at country's progress through carefully identifying metrics related to the development and then measuring them. In this study, five main socio-economic indices were chosen due to their importance in measuring the development from different perspectives including economic (GDP), technical (IDI), human wellbeing (HDI), environmental (HPI), and fertility rate perspectives. We used the World Bank data, UNDP, the ITU report, and Happy Planet Index sites to retrieve the 2015 data of GDP and fertility rate, HDI, IDI, and HPI, respectively. The full data set is shown in Table 3. 2015 data were used in this study because 2019 (or more recent) data were not available for all indices at the time this research was conducted.

Firstly, the Gross Domestic Product (GDP) per capita is a standard means of measuring well-being, living standards, and the growth of the economy [12]. It can also distinguish whether a country is developed, developing, or underdeveloped, as it indicates the impact of economic policies on the quality of life.

Next, the ICT Development Index (IDI) is an index published by the United Nations International Telecommunication Union based on internationally agreed information and communication technologies (ICT) indicators. The IDI is a standard tool that governments, operators, development agencies, researchers, and others can use to measure the digital divide and compare ICT performance within and across countries [10]. The ICT Development Index is based on 11 ICT indicators, grouped in three clusters: access, use, and skills.

Meanwhile, the Human Development Index (HDI) was developed by the United Nations as a metric to assess the social and economic development levels of countries. The HDI emphasizes the people's capabilities, not economic growth alone, as the ultimate criterion to assess a country's development. HDI is a summary measure of average achievement in life expectancy, literacy, education, and standards of living [11]. Then, the Fertility Rate (FR) is the number of children born by a woman over the course of her life in a given country [12]. It is important for economic growth, cultural stability, and more. Very low fertility rates could lead to population declines, which could be bad for the economy. Meanwhile, very high fertility rates could threaten development and stability.

Last but not the least, the Happy Planet Index (HPI) is a completely new index of human wellbeing and environmental impact [18]. HPI indicates "how well nations are doing at achieving long, happy, sustainable lives" [18]. It was introduced by the New Economics Foundation in 2006 to challenge well-established development indices such as GDP and HDI, which do not take sustainability into account. For instance, GDP is considered inappropriate, since the main aim of most people is not to be rich but to be happy and healthy. Unfortunately, Brunei HPI for 2015 is not available. Therefore, we have estimated Brunei HPI 2015 subjectively by looking at population, GDP, HDI, life expectancy, total forest area, and percentage of land area (forest). Based on the HPI scale, Brunei's HPI should be more than 44.6 because of its high GDP, high HDI, high life expectancy, and small population as compared to the number of trees, which is forest covering 52.8% of the area. Perhaps it is slightly higher than Malaysia at 55.

Table 3. Internet Performance and Socio-Economic Development indices data in 2015

Country	IP	GDP	IDI	HDI	Fertility rate	HPI
Brunei	801.712	30,967.90	5.250	0.865	1.90	**55.00***
Cambodia	829.689	1,163.20	2.780	0.563	2.60	42.34
Indonesia	480.078	3,336.10	3.630	0.689	2.40	58.92
Laos	313.631	2,159.40	2.210	0.586	2.90	57.34
Malaysia	887.92	9,643.60	5.640	0.789	1.90	54.05
Myanmar	526.133	1,194.60	1.950	0.556	2.20	51.23
Philippines	525.512	2,878.30	3.970	0.682	2.90	59.02
Singapore	1656.098	53,629.70	7.880	0.925	1.20	48.24
Thailand	1034.987	5,814.90	5.050	0.740	1.50	50.90
Vietnam	724.348	2,107.00	4.020	0.683	2.00	66.52

Note: Data were obtained from World Bank data (for GDP and fertility rate), UNDP (for HDI), the ITU report (for IDI), and Happy Planet Index (for HPI) sites.

*This figure was estimated for this research since it was unavailable

4.0 INTERNET PERFORMANCE

The Packet Loss metric provides a good indication whether part of the path is congested. Packet loss is typically caused by network congestion that in turn causes queues (e.g., in routers and switches) to fill and packets to be dropped. Packet loss is also caused by noise in the links or bit errors in the network devices [17]. The link quality can roughly be characterized by the packet loss level as follows: Less than 0.1% packet loss is excellent; between 0.1 to 1% packet loss is good; packet loss between 1% to 2.5% is acceptable; packet loss between 2.5 to 5% is poor; packet loss between 5% to 12% is very poor, and finally packet loss of more than 12% is considered bad [17]. For instance, packet loss in interactive applications such as VoIP should never exceed 1%, which basically means one voice skip every three minutes. Packet loss figures greater than 3% can significantly impact the performance of connection-oriented transports [14]. In fact, a packet loss greater than 10 – 12% is an unacceptable level of back-to-back packet loss causing long retry time outs, TCP connections begin to break and real time voice and video are not possible [17].

Figure 3 shows the packet loss scenarios as seen from SLAC to hosts in the ASEAN countries. It is shown on a log scale to accommodate a wide range of losses. It is seen the losses are falling with time. Singapore has consistently the lowest loss. This is followed by Brunei, Thailand and Malaysia in the last two years. The highest losses have been observed for Laos, Indonesia and Myanmar. In 2009, 2010, and 2019 Laos had the highest loss. This has a negative impact on interactive applications such as video conferencing and audio chat that require low packet losses. In addition, bulk data transfer applications are also affected by high packet loss because losses cause long delays for time outs also leading to high jitter.

The Round-Trip Time (RTT) metric is another indicator of the link quality. Unlike packet loss, however, whereas it is possible to reduce losses to zero, it is impossible to reduce the RTT to less than the time taken for packets to travel the distance along a fiber or copper cable or wireless links [17]. In addition to the cable or wireless delays, the measured RTT is the time taken for the packet to be accepted by router interfaces, the delay caused by queuing, and the time taken for the packet to be passed to outbound interfaces [17]. In practice for wide area network connections, the most important factors are the fiber or copper cable or wireless delays. To first order the minimum RTT indicates the route length taken by the packets, the number of hops as well as the line speeds. The RTT distribution also indicates the congestion along the path and step changes in minimum RTT can imply a route change [14].

In Figure 4, during the period of 2000 to 2002, the connection from SLAC to Indonesia utilized a geostationary satellite with a minimum RTT of over 500 msec. Since 2004, the minimum RTT is constrained by the terrestrial path and drifts between 190 to 210 milliseconds with Singapore having the lowest value (roughly 190 milliseconds) since it is generally the point of entry from SLAC for most ASEAN countries. Thailand appears to have the largest minimum RTT of about 201 msec. This is since the typical route from SLAC goes from the US to Honk Kong or Japan, then Singapore and finally to Thailand.

The Throughput metric is derived from RTT and loss and is the estimated rate at which data is delivered between pairs of hosts. From Figure 5, it can be seen that though variable from year to year, there is roughly an order of magnitude improvement so far this century. The improvement appears to be roughly exponential until around 2009, after which the improvement per year is flatter. Singapore has had the highest throughput. This is followed by the group of countries comprising Malaysia, Thailand, and Philippine. The lowest performing countries tend to be Laos, Myanmar, and Brunei.

From the observations over the period 2000-2019, a pattern can be seen. The countries in the ASEAN sub-region seem to be categorized into three groups according to IP, namely; Group 1: Singapore, Malaysia, and Thailand; Group 2: Philippines, Indonesia, and Brunei; and Group 3: Laos, Cambodia, Vietnam, and Myanmar. The countries in each group have similar Internet performance characteristics for throughput, delay, and packet loss (Figures 3 to 5). Group 1 countries are the most developed in the sub-region and have the best Internet performance, this is followed by Group 2 countries. Group 3 countries are the least developed and have the lowest Internet performance in the ASEAN sub-region.

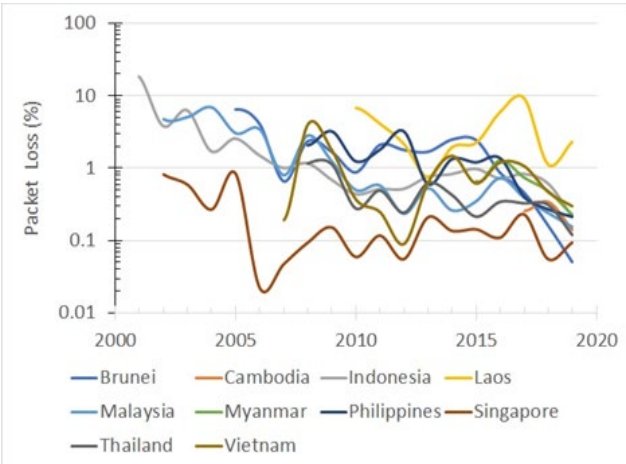


Figure 3. Percent Packet loss observed from SLAC to ASEAN countries since 1999.

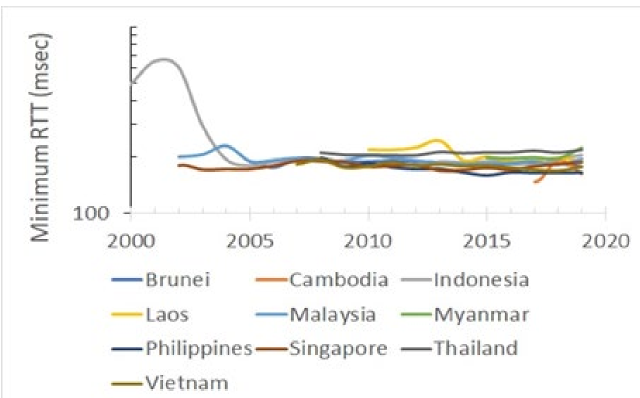


Figure 4. Minimum RTT observed from SLAC to ASEAN countries since 2000**.**

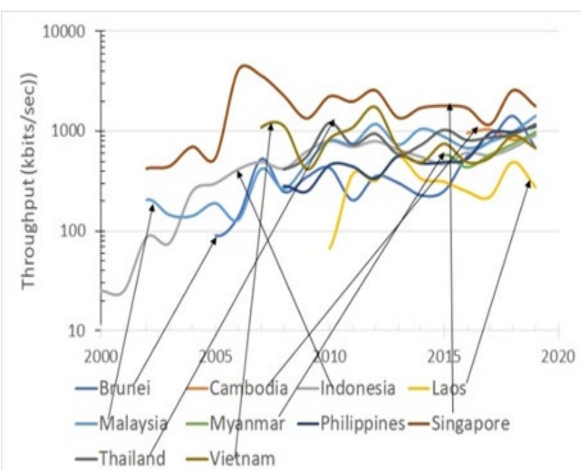


Figure 5. TCP Throughput observed from SLAC to ASEAN countries since 2000.

5.0 IMPLICATION OF SOCIAL DEVELOPMENT INDICES ON INTERNET PERFORMANCE IN ASEAN COUNTRIES

In this study, Internet performance (IP) is represented by the Throughput which reflects the actual amount of data the user can successfully transmit per second. The next phase of the analysis is to produce a initial model to examine the impact of social economic development indices on IP in ASEAN countries. To achieve this aim, the analysis began with an Exploratory Data Analysis (EDA) approach to identify systematic relations between IP and selected development indices. After establishing the suitability of the data, the assumption of normality was tested using Kolmogorov-Smirnov and Shapiro-Wilk tests to proceed with the appropriate correlation test. A parametric correlation test would be performed if the data passed the normality test, and a non-parametric correlation test would be performed otherwise. Thus, from the results of the normality tests, the GDP was the only index that was not normally distributed, and thus required the Spearman Correlation test, while the other indices were tested using the Pearson Correlation test. After establishing the correlational relationships, all the significant associated indices were then subjected to multiple regression analysis to formulate an initial model. This model was then verified by inspecting the Normal P-P Plot of Regression Standardized Residual. The use of a statistical software package, namely SPSS Statistics by IBM, was employed to perform aforementioned analytical procedures.

Exploratory Data Analysis (EDA): Exploratory Data Analysis (EDA) approach was used to identify systematic relations between the Internet performance on one hand and development indices on the other hand. A graphical scatter plot was used to observe the relationship between the Internet Performance (IP) versus GDP, IP versus IDI, IP versus HDI, IP versus FR, and IP versus HPI, as shown in Figures 6 to 10 respectively. This was also done to satisfy some assumptions for analyzing the data further.

In this study, throughput was correlated with GDP per capita as represented in Figure 6. The results revealed a strong positive correlation (R2 about 56%), implying that the stronger the GDP/Capita, the better the Internet performance. Strong economies have the potential of growing the Internet and improving its performance (throughput). Singapore has the best GDP per capita in the sub-region and very good Internet throughput. Malaysia and Thailand's GDP per capita are not far from the sub-regional average with a close to average throughput. However, Laos, Vietnam and Myanmar have low GDP per capita and corresponding low throughput.

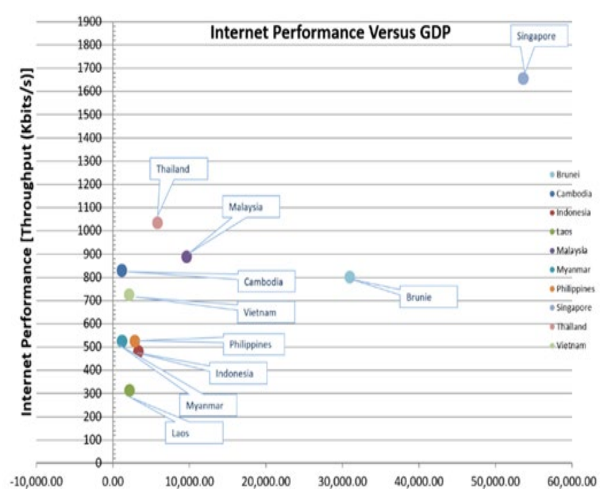


Figure 6. Internet Performance versus GDP

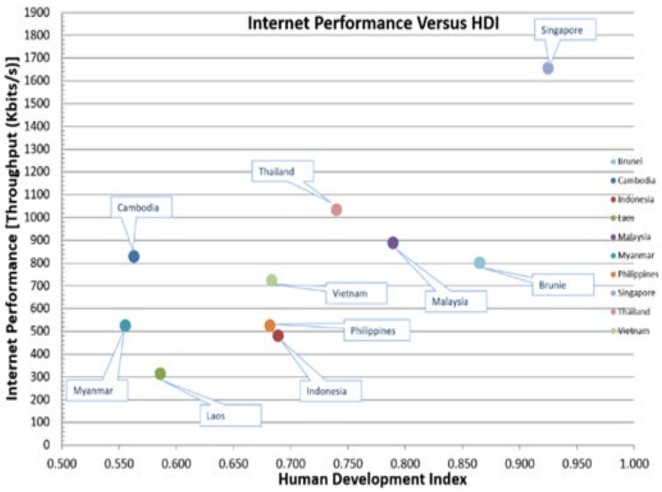


Figure 7. Internet Performance versus HDI

The HDI metric emphasizes the people's capabilities, not economic growth alone, as the ultimate criterion to assess a country's development. In Figure 7, throughput was correlated with HDI. The correlation shows that there is positive relationship between HDI and throughput, however, the correlation is a medium one (R2 about 51%). The result implies that as the HDI in the sub-region increases, the throughput also increases. This indicates that the better the life expectancy, literacy, education, and living standard in ASEAN countries, the better their countries' Internets perform. Countries like Laos, Myanmar, and Cambodia with very low HDI also have the worst Internet. On the other hand, countries like Singapore, Brunei, Thailand and Malaysia with higher HDI also have higher Internet performance. Furthermore, apart from those four countries, all the other ASEAN countries fell below the sub-regional average.

The IDI metric is a measure ICT development in a country [10]. This index was correlated with throughput in Figure 8. It is expected that a country's growth of IDI will lead to a corresponding improvement in the country's throughput. There is a strong (R2 = 71%) positive linear correlation between IDI and throughput, indicating that in ASEAN countries, on average, the higher the IDI, the better the Internet performance (see Figure 8). Laos, Cambodia and Myanmar have the lowest IDI and the poorest throughput in the sub-region while Brunei, Thailand, Malaysia and Singapore with higher IDI, also have higher throughput. Singapore has the highest IDI as well as the highest Internet performance in the ASEAN sub-region. Looking at the sub-regional average, only Brunei, Thailand, Malaysia and Singapore are above the average, while all other countries are below the sub-regional average.

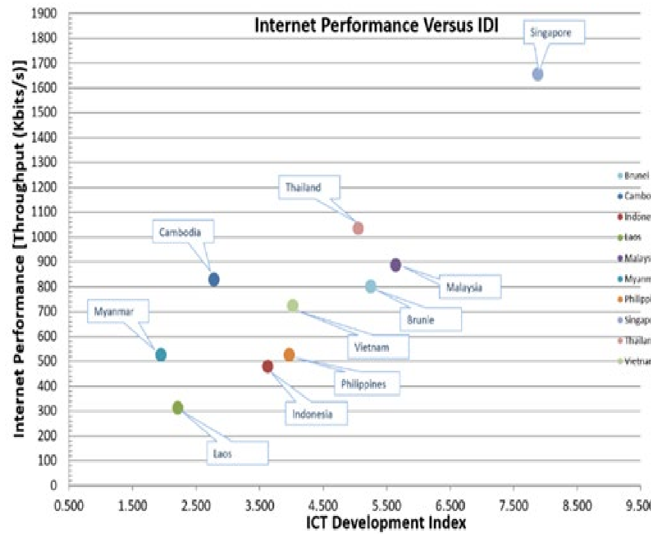


Figure 8. Internet Performance versus IDI

The Fertility Rate is the number of children born by a woman in a given country [12]. An increase in fertility rate could threaten development and stability. In Figure 9, throughput was tested with the FR to observe how ASEAN countries' Internet performances fare when associated with the fertility rate. It is apparent that there is a negative correlation between throughput and fertility. This implies that on average, the higher the fertility rate, the lower the throughput in the sub-region. Countries with more children born per woman have lower Internet performance. In addition, the fertility rate contributes only about 41% of the variability in throughput (a weak correlation). A closer look at ASEAN countries reveals that almost all ASEAN countries are below the sub-regional average for throughput, except Singapore, Malaysia, and Thailand. The lower throughput values are disturbing since achieving significant fertility decline requires education and easy access to information, and this in turn is enabled by good Internet access. Thus, countries such as Laos, Cambodia, and Philippines with high FR and low IP are particularly at risk.

The Happy Planet Index (HPI) is a completely new index of human wellbeing and environmental impact [18]. Throughput was correlated with HPI and reveals a negative linear relationship between HPI and throughput (see Figure 10). This implies that in the sub-region, the lesser the ecological efficiency with which human wellbeing is delivered, the better the Internet throughput. However, the relationship of HPI and throughput and impact of HPI on throughput is very weak in the sub-region (R2 about 16%). Vietnam has the highest HPI and a low throughput while Singapore has a low HPI and the highest throughput. Only Thailand seems to be above the sub-regional average. It should be noted that there was no HPI data for Brunei.

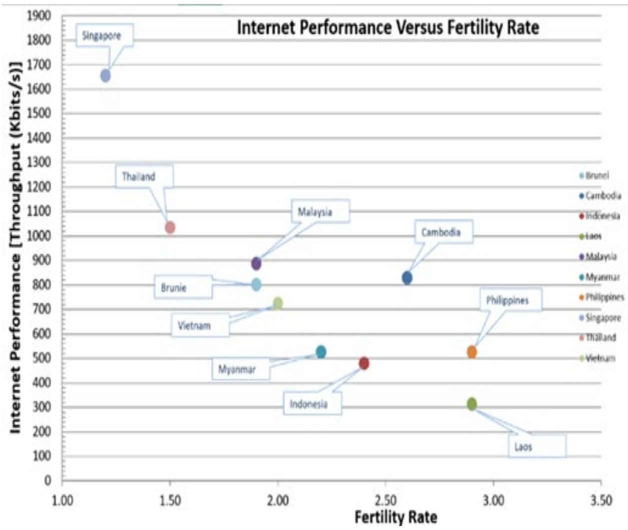


Figure 9. Internet Performance versus fertility rate

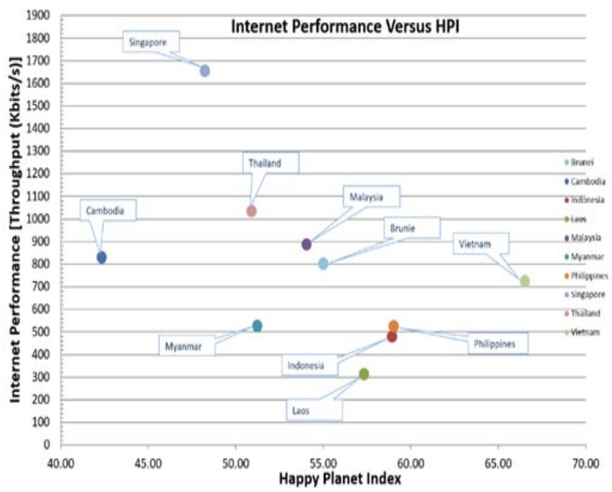


Figure 10. Internet Performance versus HPI

Normality Test: The normality test was performed using Kolmogorov-Smirnov and Shapiro-Wilk tests, as shown in table 4 and table 5 respectively. For both tests, the null hypothesis (H0) is that the tested data is normal, so the hypothesis (H1) is considered not normal, at a significant level of α=0.05. Table 4 shows the result of the Kolmogorov-Smirnov test, which revealed that the null hypothesis is rejected for the GDP indicator thus it is not normally distributed. Meanwhile, the rest of the indices had their associated null hypothesis accepted, the same goes true with regard to Shapiro-Wilk tests, which indicates that they are all normal and thus will be subjected to the Pearson Correlation test. The GDP indicator will be tested using Spearman Correlation test.

Table 4. Results of the KOLMOGOROV-SMIRNOV test.

Index	Statistic	df	Sig.	H0	Result
IP	.186	10	.200*	Accepted	Normal
GDP	.338	10	.002	Rejected	Not Normal
ICI	.148	10	.200*	Accepted	Normal
HDI	.160	10	.200*	Accepted	Normal
FR	.129	10	.200*	Accepted	Normal
HPI	.143	10	.200*	Accepted	Normal

Note: *This is a lower bound of the true significance

Table 5. Results for the SHAPIRO-WILK test

Index	Statistic	df	Sig.	H0	Result
IP	.891	10	.174	Accepted	Normal
GDP	.648	10	.000	Rejected	Not Normal
ICI	.950	10	.663	Accepted	Normal
HDI	.936	10	.506	Accepted	Normal
FR	.956	10	.744	Accepted	Normal
HPI	.984	10	.983	Accepted	Normal

Correlation Test: The associated correlation tests were performed on the data, where the Spearman Correlation test was employed to investigate the relationship between IP and HPI, while Pearson product-moment correlation coefficient analysis was run to observe the relationship between IP and the other indices. However, several assumptions needed to be satisfied in order to be able to make meaning out of the results. The data was observed to be continuous data, which satisfy the requirement of the data type interval or ratio level. Next, through visual inspection during the EDA stage of analysis, it was observed that there is a linear relationship between the IP and different indices. As for the outlier assumption, the SPSS Statistics software was set to screen and exclude outliers from the analysis. Finally, from the results of the normality analysis, it was confirmed that the data are normal, and thus suitable for further Pearson Correlation analysis.

From Table 6, it can be observed that there are relationships between IP and different indices of varying degrees. As expected, for the relationship between IP and IDI, there is a strong positive correlation, with r = 0.867, N = 10, and p = 0.001. This means that the Internet performance of a country would increase with greater advancement in ICT development, as reflected in the increasing IDI.

Table 6. PEARSON correlation results.

		IDI	HDI	FR	HPI
IP	Pearson Correlation (r)	.867**	.736*	-.848**	-.469
	Sig. (2-tailed) (p)	.001	.015	.002	.172
	N	10	10	10	10

Similarly, the relationship between IP and HDI was observed to be also significant, but at a lesser more moderate level, with r = 0.736, N = 10, and p = 0.015. This shows that Internet performance will also increase with the enhancement of human development, as shown by the increasing HDI.

Meanwhile, the other two indices had a negative relationship with IP, with FR having a strong relationship (with r = -0.848, N = 10, and p = 0.002) and HPI having a weak relationship (with r = -0.469, N = 10, and p = 0.172). Both indices indicate an inverse relationship with Internet performance, meaning that as the related indices decrease, the Internet performance would increase, with the FR having a significantly stronger effect than HPI.

As for the relationship between IP and GDP as shown in Table 7, the Spearman correlation test revealed the relationship to be weak, with r = 0.479, N = 10, and p = 0.162.

Table 7. SPEARMAN correlation result

		GDP
IP	Correlation Coefficient (r)	.479
	Sig. (2-tailed) (p)	.162
	N	10

The Proposed Regression Model: After establishing the link between IP and the indices, the cause-effect relationship was examined using Multiple Regression analysis. As with the correlation tests, similar and extra assumptions, such as homoscedasticity of data and normality of residual errors, were examined and deemed acceptable.

Table 8. The model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.992a	.985	.966	70.135083

a. Predictors: (Constant), HPI, HDI, FR, GDP, IDI
b. Dependent Variable: IP

There were several outputs from the multiple linear regression analysis stage, which includes Table 8 that shows the output for model summary, Table 9 with the analysis of variance (ANOVA) results, and finally Table 10 with the coefficient values.

The multiple linear regression was performed specifically to predict the Internet performance based on five development predictors, namely GDP, HDI, FR, IDI, and HPI. From Table 9, a significant regression equation was revealed ($F(5,4) = 51.523$, $p < 0.001$) [20], with an R^2 value of 0.985 (from Table 8). This indicates that the regression model explains 98.5% of the variance in the data.

Table 9. ANOVA Output

Model	Sum of Squares	df	Mean Square	F	Sig. (p)
Regression	1267200.912	5	253440.182	51.523	.001b
Residual	19675.720	4	4918.930		
Total	1286876.632	9			

According to Table 10, four of the five predictors (GDP, IDI, HDI, and FR) were statistically significant predictors of IP, while HPI was borderline. After further inspection of the data and validation, HPI was included in the model since its exclusion had caused the overall model to become not valid. Finally, the proposed model to show cause effect relationship is presented as:

$$IP=2724.288+0.008*GDP+275.904*IDI-3238.298*HDI-220.264*FR-8.084*HPI$$

This model translates as: Internet performance will increase with increases in GDP and IDI by a factor of 0.008 and 275.904 respectively, and decreases in HDI, FR, and HPI by a factor of 3238.298, 220.264, and 8.084 respectively.

We observed strong correlations obtained when we correlate the normalized throughput with technology or Internet related development indices, specifically with the GDP and IDI. There is a moderate to strong correlation (R^2 about 0.51) seen when we correlate the throughput versus HDI, while a negative correlation is seen when we compare the throughout versus HPI and fertility rate. Table III shows the R^2 for the correlations of PingER measurement with GDP, IDI, HDI, HPI and Fertility Rate.

From the regression model, it illustrates the causal-relationship between Internet performance and the five predictors of development. The model shows that enhanced Internet performance can be achieved when GDP and IDI increases, while at the same time HDI, FR, and HPI decreases, by certain respective amounts.

Regression Validation of the Proposed Model: According to [19], the final step of this investigation into the relationship between IP and the development indices requires a regression validation step of the proposed model. This was done to verify the model in order for it to be useful and applicable for future studies in or similar to this field.

The validation was done through several procedures, namely validation using R^2 and analysis of residuals. Initial readings showed $R^2=0.985$, which is very high. It is then recommended that further analysis is needed, since a high reading of R^2 would indicate, but not confirm a good fit. Thus, the next step in the verification of the model was through residual analysis. For the purpose of this research, which was exploratory with nature and characteristically restricted according to the availability of limited data, the graphical analysis of residuals was employed. This was because even though this technique is not quantitatively precise, it has an advantage over numerical methods in that it readily illustrates a broad range of complex aspects of the relationship between the data and the model.

According to [19], a residual can be defined as the difference between the actual data and the calculated result using the regression model equation. Mathematically, the definition of the residual for the ith observation in the data set is written as:

$$e_i = y_i - f(x_i; \beta) \quad (2)$$

where y_i denotes the ith response in the data set, while x_i is the vector of explanatory variables (development indices). The numerical results of each observation were then drawn in a scatterplot. It was recommended [19] that if the residuals behaved randomly, in other words there is no obvious trend observed in the visual inspection of the scatter plot, then it suggests that the model fits the data well. From a visual inspection of Figure 11, it can be concluded that the model developed in this study is a good fit to the data, which supports the relationship between IP and the development indices.

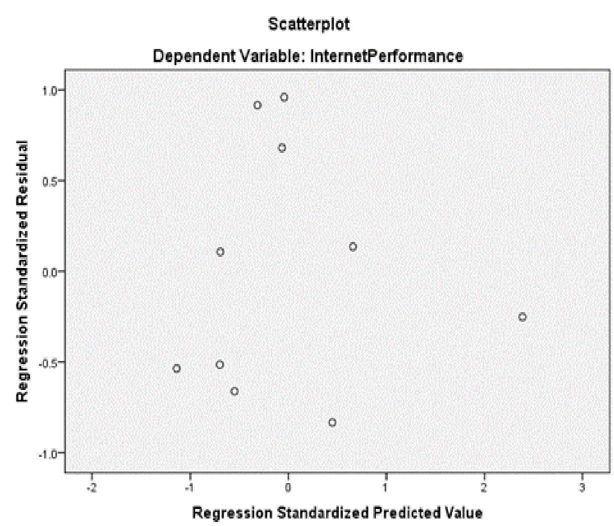


Figure 11. Residual scatterplot

6.0 CONCLUSION

This paper examines the Internet performance in ASEAN countries and analyzes the correlation between Internet performance, specifically the average throughput and several socio-economic development indices namely, GDP, IDI, HDI, FR, and HPI. A novel model was designed and validated to study the effect of socio-economic development indices on the Internet performance. A real end-to-end performance measurement from SLAC to ASEAN countries using PingER was performed to test the Internet performance 2000-2019. Meanwhile, the GDP and Fertility ratio data were obtained from the World Bank. The HDI and IDI are obtained from the UNDP Human Development Reports and the ITU Reports respectively.

Our study discovers several interesting findings. Firstly, the ASEAN countries seem to be categorized into three groups according to Internet performance. Further, the development level in ASEAN countries is directly proportional to Internet performance. Secondly, improving GDP and IDI by a factor of 0.008 and 275.904 respectively will increase the Internet performance. Meanwhile, the Internet performance will increase with decreases in the non-technological indices HDI, FR, and HPI by a factor of 3238.298, 220.264, and 8.084 respectively. Finally, although we do not suggest that PingER data replace socio-economic indices, the PingER data can be considered as a useful complement, specifically for monitoring ICT related development indices as well as for highlighting inquisitive inconsistencies between Internet performance metrics and development indices.

The result of the study followed and observed a trend in the relationship between the socioeconomic development indices and internet performance in ASEAN countries up to 2015. The data of the covid-19 pandemic era (2020-2022) were not included. This data, though important is a spike data that may bias the model as it is a sharp seasonal variation. It will not indicate a normal trend in internet performance viz a viz the socioeconomic development indices. Our study is interested in developing a model that will serve the post-covid-19 era that will model the normal trend and not a seasonal variation that is occasioned by spike data. Furthermore, these socioeconomic development indices are collected by several international organizations, and they took a long time and hard process to collect these data. Even though internet performance data is available up to date, these socioeconomic development indices data are not available for all metrics for all ASEAN countries. The model is still relevant to the current situation (post-covid-19). This is because the model is meant to model and address normal variations and not seasonal variations. With the post-covid-19, the spike in internet use has normalized as usage of the internet had returned to normal. The sharp spike in the covid-19 pandemic data would have introduced seasonal variation bias into the model and confused the model to predict well now that the sharp spike in internet use is no more there in internet use.

Table 10. Regression coefficients values

Coefficients [a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2724.288	371.755		7.328	.002
	GDP	.008	.003	.364	2.570	.062
	IDI	275.904	50.814	1.307	5.430	.006
	HDI	-3238.298	821.760	-1.068	-3.941	.017
	FR	-220.264	69.352	-.329	-3.176	.034
	HPI	-8.084	4.561	-.143	-1.772	.151

a. Dependent Variable: IP

Future work: this research effort is not without its apparent limitations; therefore, it is recommended that further discussion on the complexities and improvements to the model can be made by finding and including more up-to-date data for socio-economic development indices. Further, the use of ARCH or GARCH models or other Simultaneous Models can be considered to cater for outliers

Acknowledgement

We would like to thank Universiti Utara Malaysia for supporting this study through the Special Research Grant (SO Code: 14330). Special thanks to Assoc. Prof. Dr. Suzilah bt. Ismail for her practical advice on data analysis and numerous discussions. We are also grateful for the valuable comments and suggestions from the reviewers that have helped to improve the quality of our paper.

References

[1] Internet Society. 2015. Unleashing the Potential of the Internet for ASEAN Economies. The Internet Society (ISOC) and TRPC Pte Ltd.

[2] Gurría, A. 2015. Towards Inclusive and Sustainable Growth in the ASEAN Economic Community. Davos World Economic Forum, Switzerland.

[3] Wang, T.-Y. and Chien, S.-C. 2007. The Influences of Technology Development on Economic Performance: The Example of ASEAN Countries. *Technovation*. 27(8): 471–488.

[4] Srinuan, C., Rohman, I.K., Srinuan, P. and Bohlin, E. 2010. Digital Divide in ASEAN Countries: Explaining the Gap. *Competitiveness of the ASEAN Countries: Corporate and Regulatory Drivers*. London: Edward Elgar. 153–172.

[5] Ourworldindata.org. 2017. How Many Internet Users Does Each Country Have? Available:http://www.internetlivestats.com/internet-users-by-country/

[6] Badran, M. F. and Badran, H. F. 2011. Promoting Economic Growth by Broadband Development in Emerging Countries: An Empirical Study. *2011 Technical Symposium at ITU Telecom World (ITU WT)*, Geneva. 199-204.

[7] Les Cottrell, R. PingER - Internet End-to-End Performance Monitoring - SLAC Confluence. https://confluence.slac.stanford.edu/display/IEPM/PingER.

[8] Whittaker, Z. Syria suffers Internet 'blackout'; cut off from the outside world. http://www.zdnet.com/article/syria-suffers-internet-blackout-cut-off-from-the-outside-world/

[9] Abdullah, J., Lenando, H. and Jambli, M.N. 2013. Towards Achieving Intelligent Event Correlation for End-to-end Internet Performance Monitoring. *2013 IEEE Conference on Open Systems (ICOS)*, Kuching, 127-132.

[10] ITU. 2015. Measuring the Information Society Report. *International Telecommunication Union. 2015.* Measuring the Information Society Report. ITU Publications

[11] Undp.org. 2015. United Nations Development Programme: Human Development Reports. http:// www.hdr.undp.org/.

[12] WorldBank. 2018. World Development Indicators| World DataBank," The World Bank. http://databank.worldbank.org/data.

[13] Barford, P. et al. 2002. A Comparison of Active and Passive Methods for Measuring Packet Loss. *University of Wisconsin-Madison Technical Report*.

[14] Pezaros, D. 2005. *Network Traffic Measurement for the Next Generation Internet*., PhD, Lancaster University.

[15] Mathis, M., Semke, J. and Mahdavi, J. 1997. The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm. *ACM SIGCOMM Computer Communication Review*. 27(3): 67–82.

[16] Jensen, M. 2003. Benchmarking Connectivity in Africa: Towards Effective Indicators for Measuring Progress in the Use of ICTS in Developing Countries. *Joint NECE/ UNCTAD/ UNESCO/ ITU/ OECD/ Eurostat Statistical Workshop: Monitoring the Information Society: Data, Measurement and Methods*, Geneva.

[17] Zennaro, M. et al. 2006. Scientific Measure of Africa's Connectivity. *Information Technologies & International Development.* 3(1): 55–64.

[18] HPI. 2016. Happy Planet Index. http://happyplanetindex.org/.

[19] Kmenta, J. 1986. Elements of Econometrics (2nd ed.). New York: MacMillan.

[20] Wikipedia. 2018. F-test. https://en.wikipedia.org/wiki/F-test. 25/Oct/2021