Clock Skew Based Computer Identification on Different Types of Area Networks

Nola Verli Herlian[1], Komang Oka Saputra[2], I.G.A.K Diafari Djuni Hartawan[3]

[1,2,3]Department of Electrical Engineering, Faculty of Engineering,
Udayana University, Bali – Indonesia
[1]nola.verli@gmail.com

Abstract—The increase of client devices along with the growth of internet access currently affects to security threats at the user's identity. Identifiers that commonly used today, such as SSID, IP address, MAC address, cookies, and session IDs have a weakness, which is easy to duplicate. Computer identification based on clock skew is an identification method that is not easily duplicated because it is based on the hardware characteristics of the device. Clock skew is the deviation of the clock to the true time which causes each clock to run at a slightly different speed. This study aims to determine the effect of network types to the clock skew stability as a reliable device identification method. This research was conducted on five client computers which running windows and linux operating systems. The measurement was conducted based on three different types of area networks, i.e., LAN, MAN, and WAN. The skew estimation was done using two linear methods i.e., linear programming and linear regression. The measurement results show that the most stable clock skew is found on the LAN measurement because it meets the threshold tolerance limit i.e., ±1 ppm. Skew estimation using linear programming method has better accuracy than linear regression method.

Index Terms—**Clock Skew, Computer Identification, Linear Programming, Linear Regression**

I. INTRODUCTION

Nowadays, internet access is growing rapidly and easy to reach. The number of client devices is ever-increasing. However, when the device is connected to a public access point network, there are threats from the security side against the client's identity. Therefore, identification of user devices becomes the important thing. Device identification is used to detect access from unauthorized account to the internet, for example through the password theft. Identifiers that commonly used today, e.g., SSID (Service Set Identifier), IP address, MAC address, cookies, or session ID have a weakness, which is easy to duplicated [1]-[3]. This research was done on one of the identification methods that is not easily duplicated i.e., clock skew based device identification [4].

Clock skew is defined as a clock deviation from the true time [5]. In every physical clock, the clock signal is generated by a crystal oscillator. Clock skew is an inescapable phenomenon that causes each clock to tick at a slightly different speed. This is caused by mechanical accuracy imperfections of the crystal oscillator during the manufacturing process. Even though, for crystals which have the same series, type and production date [6]. Clock skew is usually expressed in microseconds per second

($\mu s/s$) or equivalent to parts per million (ppm) [4] which means that there is one microsecond deviation in every second. Clock skew is a reliable device identification method because it based upon the hardware characteristics.

In the clock-skew based device identification, the server computer (fingerprinter) periodically collects timestamps from the client computer (fingerprintee). Based on these data, the clock skew graph which consists of the test duration and offset is obtained. The offset is difference of the time between two clocks [7], which is the time when data is sent from the client computer and the time when data is received by the server computer.

In previous research, clock skew-based remote device identification has been successfully proven [4], [8]-[11]. This research observed about the performance of clock-skew based computer identification over LAN (Local Area Network), MAN (Metropolitan Area Networks), and WAN (Wide Area Networks) in the Udayana University area.

II. LITERATURE REVIEW

A. Clock

A clock is a device that ticks in regular frequency. Clock consists of two main components i.e., oscillator and counter. An oscillator is an electronic component that made of piezoelectric materials, such as quartz crystals or ceramic resonators to generate electrical signals at regular interval

with accurate frequency. Meanwhile, counter counts the number of the ticks that generated by the oscillator. [12]

The clock value at any given time (t) is denoted by $C(t)$. The value of $C(t)$ increases discretely, but it can be estimated by a derivative function $C'(t) = dC(t)/dt$. Some factors, e.g., manufacturing deficiency, temperature changes, aging, power supply variations cause the oscillator have a deviation from its nominal frequency. Thus, the actual clock rate corresponds to this equation $C'(t) = 1 + \rho(t)$, where $\rho(t)$ is a drift of the clock. [12]

B. Clock Skew

On the terminology of clock skew, there are four most important terms i.e., offset, frequency, skew, and drift. Clock offset is the time difference between the time that shown by client computer (C_c) and server computer (C_c). Clock frequency is the tick rate at which the clock running on. It is symbolized as $C_c'(t)$ and $C_s'(t)$. Clock skew is the difference of frequency rate between two clocks which expressed by $(t) = C_c'(t) - C_s'(t)$. Clock skew is the first derivative of offset to the time. Meanwhile, drift is the second derivative of offset to the time $(C_c''(t) - C_s''(t))$. [13]

Clock skew based computer identification is performed by comparing the frequency between two clocks, where one of them is assumed to represent the true time. Clock skew can be calculated by analyzing the drift of the client clock (C_c) from the server clock or reference clock (C_s). Since, a stable reference clock is usually unavailable, the possible process to get a clock skew value is by accumulate the drift of a observed clock from another clock. Therefore, the clock skew value for a node reflects a drift shift from the observed clock to the reference clock. The specified skew value from each clock is unique and depends on the hardware characteristics. [5]

C. Clock-Skew Estimation Methods

In this research, clock skew value was estimated using two different methods i.e., linear regression and linear programming. Both of these methods have often been used in previous studies. [5], [10], [14]-[16]

a. Linear Regression

Linear regression algorithm is a method that uses the average of the time on the reference clock and offset points to estimate the clock skew [5]. This algorithm is usually applied by using the Least Squares Fitting method to estimate the skew value [14]. Least squares method could find the line $y = \delta x + \psi$ from the given offset points, where δ is the line slope and ψ is intercept of y axis. The linear equation of the line is obtained by the following equation [10]:

$$\sum_{i=1}^{n} \left(d_i - \left(\delta \cdot t_i^o + \psi \right) \right)^2 \tag{1}$$

The values of slope (δ) and intercept (ψ) could be determined through the following equations:

$$\delta = \frac{n\sum X_i Y_i - \left(\sum X_i\right)\left(\sum Y_i\right)}{n\sum X_i^2 - \left(\sum X_i\right)^2} \tag{2}$$

$$\psi = \frac{\left(\sum Y_i\right)\left(\sum X_i^2\right) - \left(\sum X_i\right)\left(\sum X_i Y_i\right)}{n\sum X_i^2 - \left(\sum X_i\right)^2} = \bar{Y} - \delta\bar{X} \tag{3}$$

Linear regression algorithm has simple computation, and easy to apply. However, skew estimation using this method is strongly influenced by outliers whose values are deviate from the average. Thus, this method is inappropriate to be applied on the noisy data. [14]

b. Linear Programming

Linear programming algorithm is a method that finds the lower bound of set of data points to estimate the clock skew. From the given offset points $(t_1^o, d_1),...., (t_n^o, d_n)$, this algorithm could find the line $y = \delta x + \psi$, where δ is the line slope and ψ is intercept of y axis. The problem constraint of this method is determined as

$$\delta \cdot t_i^o + \psi \leq d_i \tag{4}$$

Linear programming should be able to minimize the following equation [10]:

$$\frac{1}{n} \sum_{i=1}^{n} \left(d_i - \left(\delta \cdot t_i^o + \psi \right) \right) \tag{5}$$

Compared to linear regression method, linear programming has better tolerance against the outliers. This method can provide a stable clock skew estimation even while data are affected by outliers. However, linear programming has more complex computation. [14]

III.   RESEARCH METHODOLOGY

This section describes about the steps to obtain the graph and clock skew estimation value, and research scenarios in different network types.

A.   *Clock-Skew Estimation Procedure*

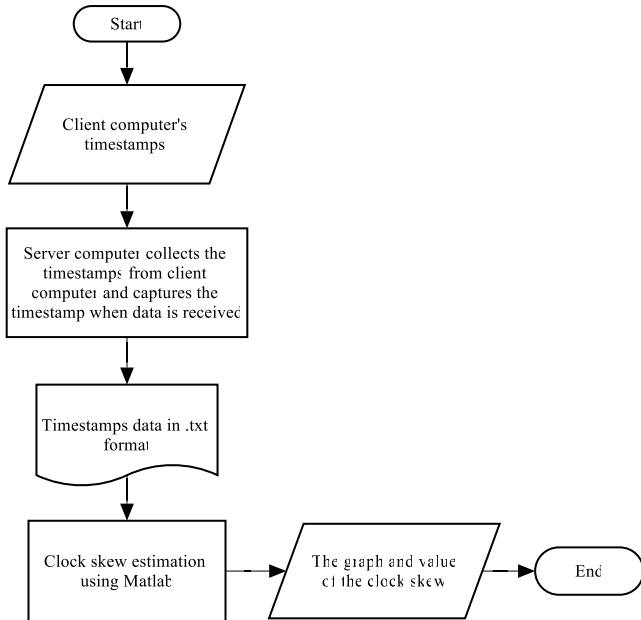The process of clock skew estimation is shown in Figure 1.

Fig. 1. Clock-Skew Estimation Procedure Flowchart

The first step is done by running the Python socket-programming program on both server computer and client computer. Then, the client computer captures its timestamp and sends it to the server computer at regular interval. The server computer also captures the timestamps when the data are received to obtain the offset value. Timestamp is sent every 0,5 seconds for 30 minutes in each measurement. These data are stored in .txt file format. In the final step, clock-skew is estimated by running linear programming program on Matlab R2015a software to obtain the graph and clock-skew value.

B.   *Research Scenarios*

This research was conducted on three different network types i.e., wireless LAN, MAN, and WAN. In addition, the measurement was also performed on wired LAN to obtain the reference value of clock skew for each client computer, that used to compare with other wireless measurements

a. LAN (Local Area Network)

The timestamps measurements on Local Area Network consist of wired and wireless measurements. On wired LAN measurement, the server computer and the client computer are connected through a UTP cable. Whereas, on wireless LAN measurement, the server computer and the

client computer are connected to the same access point. The scenarios for these measurements are shown in Figure 2 and Figure 3.
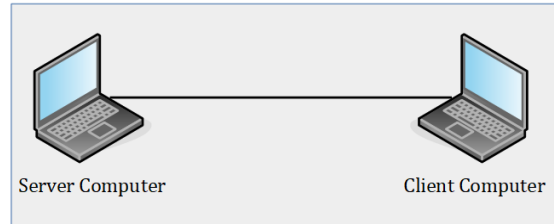

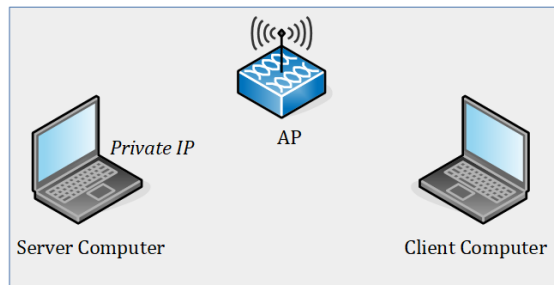Fig. 2. Research Scenario on Wired Local Area Network

Fig. 3. Research Scenario on Wireless Local Area Network

b. MAN (Metropolitan Area Network)

On the timestamps measurement over Metropolitan Area Network, the server computer and client computer are placed in different campus areas. Both computers are connected through the MAN connection on Udayana University network environment. The scenario of this measurement is shown in Figure 4.
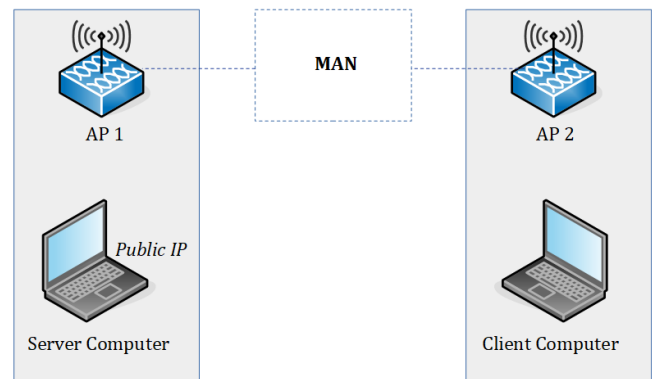

Fig. 4. Research Scenario on Metropolitan Area Network

c. WAN (Wide Area Network)

On the timestamps measurement over Wide Area Network, the server computer is placed in the campus area, whereas the client computer is placed outside the campus area. Both computers are connected through the WAN connection. The scenario of this measurement is shown in Figure 5.
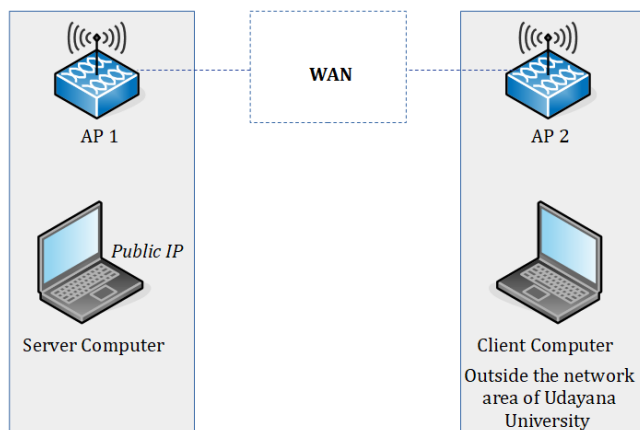
Fig. 5. Research Scenario on Metropolitan Area Network

IV.   RESULTS AND DISCUSSIONS

This experiment was conducted on five client computers, which consists of three client computers which running windows operating system and two client computers which running linux operating system. A server computer as a fingerprinter runs linux operating system. Timestamp collection from client computers was done through TCP connection with 30 minutes testing duration for each client computer.

A. *Clock Skew Estimation Results*

The skew estimation results using linear programming and linear regression methods on different types of networks at Udayana University network environment are shown in Table 1.

TABLE 1
CLOCK SKEW ESTIMATION RESULTS

Client Computer	OS	Network Type	Clock Skew (ppm)	
			LP	LR
Client 1	Windows 8.1	Wired LAN	-38,7281	-39,798
		LAN	-38,9472	-18,5570
		MAN	-8,9123	-15,2120
		WAN	-23,0726	-22,1705
Client 2	Windows 7 Ultimate 32-bit	Wired LAN	20,9857	20,9140
		LAN	21,0647	21,0670
		MAN	15,2119	22,1280
		WAN	12,3417	14,921
Client 3	Windows 7 Ultimate 64-bit	Wired LAN	-4,2471	-4,2454
		LAN	-4,3247	-8,8840
		MAN	-2,6339	-2,7005
		WAN	-0,9759	-6,9824
Client 4	Linux Mint 19 Cinnamon	Wired LAN	37,5348	37,5380
		LAN	37,5781	37,6620
		MAN	16,8083	29,2790
		WAN	4,5461	3,7852
Client 5	Linux Ubuntu 16.04	Wired LAN	0,8691	0,8704
		LAN	0,7589	2,1224
		MAN	12,9238	12,9660
		WAN	10,8324	17,7830

In this experiment, the value of clock skew varies in range of -39,798 ppm to 37,5380. The clock skew estimation result on wired LAN measurement is determined as the reference clock-skew value because it is not much affected by factors that can influence the accuracy of clock-skew estimation, such as jitter and unstable network connection. As seen in Table 1, the experiment to all different computers show the same result, that is the most stable clock skew value is performed on wireless LAN measurement. The clock-skew difference value between wireless LAN measurement and the reference clock-skew result meet the threshold tolerance limit i.e., ±1 ppm [3]. Whereas, the measurements over MAN and WAN obtain the significant difference from the reference clock-skew result.

Analysis of this research is restricted on the effect of the network environment to the clock skew stability under different network types. Thus, it does not take into account the external effect such as temperature change, workload of the computer, or Network Time Protocol (NTP). Significant differences value at the MAN and WAN measurement results are caused by factors that come from outside the effect of the network environment. These factors cause this research to have different result from the first clock-skew based fingerprinting research that was explained by T. Kohno [4] which shows that the clock skew values are stable on various distance and types of networks.

B. *The Appropriate Method*

The graphs of clock skew estimation are shown in Figure 6 and 7. Both of these graphs are the result of the clock skew measurements on Client 1 over wired and wireless LAN. The red line shows the measurement result of linear regression method, whereas the blue line shows the measurement result of linear programming method.

As seen in Figure 6 and 7, the offset set has decreases with a negative slope. It means that the frequency on the client computer's clock has lower rate than the frequency of the server computer's clock. The frequency difference between client computer and server computer will increase along with the increase of the slope.

Figure 6 shows the clock skew estimation graph which not influenced by outliers. Both of the linear method produce almost the same result. The results are -38,7281 ppm on linear programming method and -39,798 ppm on linear regression method.
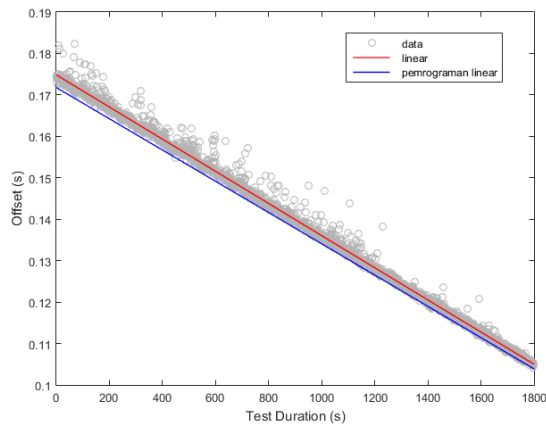
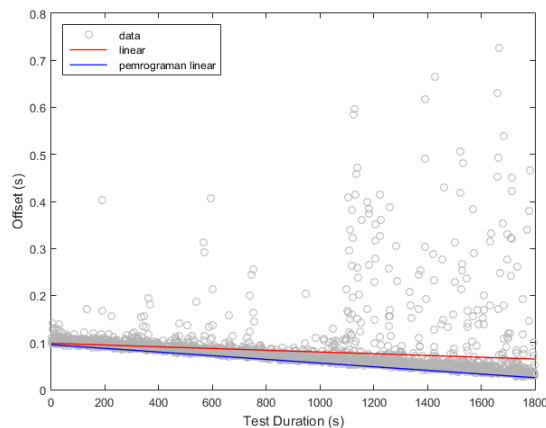Fig. 6. Clock skew estimation graph which not influenced by outliers

Fig. 7. Clock skew estimation graph which influenced by outliers

Figure 7 shows the clock skew estimation graph which influenced by outliers whose values are far from the average offset set. This causes different measurement results between both linear methods. The result of linear programming method is -38,9472, meanwhile the result of linear regression method is -18,5570. The result of linear programming method shows the closest value to the reference value.

In this experiment, most of the outliers were found above the offset set, so the appropriate method to be applied in this case is lower bound with linear programming method.

V. CONCLUSION

This research was conducted on three types of networks, i.e., LAN, MAN, and WAN. The measurement result show that the most stable clock skew value is performed on wireless LAN measurement because its difference result to the reference clock-skew result meet the threshold tolerance limit i.e., ±1 ppm [3]. The measurements over MAN and WAN obtain the significant difference from the reference clock-skew result. This is caused by factors that come from outside the effect of the network environment. Clock skew estimation using linear programming method has better

accuracy than linear regression method because it has optimal solution to solve outliers problem.

REFERENCES

[1] K. O. Saputra, W. C. Teng, and Y. H. Chu, "A clock skew replication attack detection approach utilizing the resolution of system time," IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2015, pp. 211-214.

[2] K. O. Saputra, W. C. Teng, and T. Nara "Hough transform-based clock skew measurement by dynamically locating the region of offset majority," IEICE TRANSACTIONS on Information and Systems. 2016, pp. 2100-2108.

[3] D. J. Huang, K. T. Yang, C. C. Ni, W. C Teng, T. R Hsiang, and Y. J. Lee, "Clock skew based client device identification in cloud environment," 26th IEEE International Conference on Advanced Networking and Applications, 2012, pp. 526-533.

[4] T. Kohno, A. Broido, and K. C. Claffy, "Remote physical device fingerprinting," IEEE Transactions on Dependable and Secure Computing, vol. 2, no. 2, April-June 2005, pp. 93–108.

[5] I.E. Bagci, and U. Roedig. "Node identification using clock skew," Proceedings of the 5th Workshop on Real-World Wireless Sensor Network, Berlin : Springer Verlag, 2013, pp. 111-123

[6] F. Lanze, A. Panchenko, B. Braatz, and A. Zinnen, "Clock skew based remote device fingerprinting demystified," Globecom 2012 – Communication and Information System Security Symposium, 2012, pp. 813-819.

[7] D. J. Huang, W. C. Teng, C. Y. Wang, H. Y. Huang, and J. M. Hellerstein, "Clock skew based node identification in wireless sensor networks," IEEE Global Communications Conference, 2008, pp. 1877–1881.

[8] L. Polčák and B. Franková, "On reliability of clock-skew-based remote computer identification," 11th International Conference of Security and Cryptography (SECRYPT), 2014, pp. 2291-2298

[9] L. Polčák and B. Franková, "Clock-skew-based computer identification: traps and pitfalls," Journal of Universal Computer Science, vol. 21 no. 9, 2015, pp. 1210-1233.

[10] D. J. Huang, K. T. Yang, W. C. Teng, and G. M. Chiu, "Design of client device identification by clock skew in clouds," IEEE International Conference on Automation Science and Engineering (CASE), Taipei, Taiwan, 2014, pp. 1133-1138.

[11] B. Martin, M. Tummala, , J. McEachen, "Multi-homed device detection using clock skew," Proceedings of The 11th International Conference on Signal Processing and Communication Systems (ICSPCS). Australia: IEEE, 2017.

[12] L. Gomes and J. M. Fernandes, Behavioral Modeling for Embedded Systems and Technologies: Applications for Design and Implementation, New York : Information Science Reference, 2010.

[13] S. Moon, P. Skelly, and D. Towsley, "Estimation and removal of clock skew from network delay measurements," in 18th Annual Joint Conference of the IEEE Computer and Communications Societies, vol. 1, March 1999, pp. 227–234.

[14] S. Jana and S. K. Kasera. "On fast and accurate detection of unauthorized wireless access points using clock skews", Proceedings of the 14th ACM International Conference on Mobile Computing and Networking, 2008, pp. 104-115.

[15] M. Cristea and B. Groza, "Fingerprinting smartphones remotely via ICMP timestamps," IEEE Commucations Letter, vol. 17, no. 6, 2013, pp. 1081-1083

[16] X. Mei, D. Liu, K. Sun, and D. Xu "On feasibility of fingerprinting wireless sensor nodes using physical properties," 2013 IEEE 27th International Symposium on Parallel & Distributed Processing, 2013, pp. 1112-1121.