www.jatit.org



E-ISSN: 1817-3195

# DETECTING ANOMALIES IN AOA DATA MEASURED FOR INDOOR POSITIONING TO IMPROVE ACCURACY

### HYEONGSEOL SHIN<sup>1</sup> QIANFENG LIN<sup>2</sup>, JOOYOUNG SON<sup>3</sup>

<sup>1</sup> Division of Marine System Engineering, Korea Maritime and Ocean University, Busan, Korea

<sup>2</sup>Department of Computer Engineering, Korea Maritime and Ocean University, Busan, Korea

<sup>3</sup>(Corresponding author) Division of Marine IT Engineering, Korea Maritime and Ocean University, Korea

E-mail: 1tlsgudtjf123@naver.com, 2jmulqf@outlook.com, 3mmlab@kmou.ac.kr

#### ABSTRACT

Ships are made of steel. The space of ship's environments is generally narrow. In the ship's environments noises are included in the signal. Accurate indoor positioning technology can contribute to preventing infectious diseases by identifying and locating infected people at an early stage. As a method, this paper makes the Angle of Arrival using Bluetooth 5.1 to locate infected people. The function of Anomaly Detection methods as filters may be used to eliminate outliers in elevation and azimuth angles when measuring the Angle of Arrival. Three different Anomaly Detection methods are evaluated in this paper, they are Generalized Extreme Studentized Deviation, LevelShift, and Persist. As a result, LevelShift shows an improvement by 14% in Root Mean Square Error (RMSE) and Generalized Extreme Studentized Deviation shows by 5%. Persist, however, gets worse by 2%.

Keywords: Indoor Positioning, Bluetooth 5.1, Angle of Arrival, Anomaly Detection

#### 1. INTRODUCTION

Large ships like cruise ships, and merchant ships passing through many countries may easily be media for transmitting infectious diseases in the world. Because the space of cruise ships is narrow. Cruise ships are difficult to prevent the spread of infectious diseases such as COVID-19. The location of infected people is becoming important to prevent spreading infectious diseases.

This paper suggests using an indoor positioning system (IPS) to prevent spreading infectious diseases by diagnosing the infected person in the ship. There are many IPS technology to measure user locations such as Wi-Fi, LTE, and Bluetooth. Especially, this paper uses Bluetooth Lower Energy (BLE) 5.1 to estimate user location. Therefore, if a confirmed person occurs on a ship, epidemic prevention staff can isolate the person by identifying the person's location. In this situation, IPS technology needs to be applied in the ship.

From the IPS, the signal can obtain one information, like the strength of the signal, and the distance from the signal to estimate a location. Because of the noise problem like reflection, distortion, and attenuation, various methods try to improve precision. In this paper, Angle of Arrival (AoA) is used to estimate a location. When the antenna receives the signal transmitted from the tags, the antenna measures the signal of AoA. This paper also suggests Anomaly Detection (AD) methods eliminating outliers in the data.

This paper is organized as follows. Section 2 shows previous works. Bluetooth direction finding using AoA is explained in Section 3. Section 4 describes ADs. Section 5 shows the experiment environment result of applying different AD methods. Finally, Section 6 describes the conclusion and the future study.

#### 2. PREVIOUS WORKS

AoA indoor positioning technologies can reduce multipath effects more effectively than Receiver Signal Strength Indicator. There are serious literature on AoA using Bluetooth indoor positioning algorithms. In the ref. [1], a BLE transmitter indoor positioning method is proposed based on AoA estimation. The ref. [2] evaluates experimentally, for the first time, the accuracy of an indoor positioning system based on the AoA mechanism adopted by the Bluetooth standard. The ref. [3] uses AoA method to locate in Bluetooth receiver through advertising channel. The ref. [4] proposes approach can be used for digital beamforming to increase the operating range or 31<sup>st</sup> October 2022. Vol.100. No 20 © 2022 Little Lion Scientific

#### ISSN: 1992-8645

www.jatit.org



Figure 2: measuring AoA [16]

Bluetooth direction finding uses a CTE (Constant Tone Extension) to measure AoA. Because BLE uses Gaussian frequency shift keying modulation, the frequency changes continuously through the packet to differentiate between the different symbols, or 1's and 0's, that are sent. The receiver may not be able to estimate the phase of the signal without a stable frequency. Furthermore, the frequency of the CTE is dependent upon the channel on which it is transmitted. Since the CTE is 250 kHz higher than the channel center frequency, it must be considered when measuring the distance travelled. The CTE is appended to the packet.

To determine the phase of incoming signals, Bluetooth direction finding uses IQ sampling, an inphase and quadrature sampling technology. The receiver can IQ sampling based on CTE. With IQ sampling, the receiver mixes the incoming signal with the local oscillator (LO) at 0- and 90- degrees offsets. The result is two orthogonal functions known as the in-phase (I) and quadrature (Q) components. Figure 3, these components are plotted as axis, and the magnitude and angle of the vector formed by adding the I and Q components can be used to measure the amplitude and phase of the main signal. To determine the direction of the transmitter, the receiver switches antennas as part of the AoA method.

The transmitter inserts a CTE into the packet, resulting in a fixed and constant frequency for a specified portion of the packet. Figure 4 illustrates the CTE structure. Thus, the receiver can determine the phase of the waveform by sampling its IQ components. Using this method for multiple antennas, the receiver can determine from which angle the transmitter signal originates.

decrease the level of power consumed by BLE devices in Internet of Things applications.

The ref. [5] shows Bluetooth 5.1 specification introduced the use of AoA information which enables the design of novel low-cost indoor positioning systems. But AoA is not always accurate because of noise. The ref. [6] shows AoA-based approaches are prone to noise, multi-path, and pathloss effects. The ref. [7] proposes there are several issues including Carrier Frequency Offset (CFO), multipath effect, Inter-Symbol Interference (ISI), noise, and phase shifting faced by the AoA. In fact, the noise can make AoA become inaccurate.

The ref. [8] shows incorrect identification impulses responses lead to large AoA estimation errors. The ref. [9] shows the electromagnetic interaction can make interference and distort the RFID fingerprint measurements used for AoA estimation. The ref. [10] shows outliers appeared to AoA estimation. But these papers didn't consider the patterns of outliers in AoA estimation. That could reduce accuracy of indoor positioning. The patterns of outliers may be detected by AD methods. So, this paper suggests an AD to identify the patterns of outliers. It also analyzes AoA data using the AD to identify the outliers appeared to the patterns. The result of applying the AD to AoA data shows improved accuracy of indoor positioning.

#### **3. BLUETOOTH DIRECTION FINDING**

AoA is a key principle of Bluetooth direction finding. With the use of antenna arrays, phase difference can be determined. Then, the location can be measured. Figure 1 shows what AoA method is, and Figure 2 shows how to measure AoA.



Figure 1: AoA method [16]

#### Journal of Theoretical and Applied Information Technology

<u>31<sup>st</sup> October 2022. Vol.100. No 20</u> © 2022 Little Lion Scientific



Figure 3: Unit circle for explaining IQ [15]



Figure 4: CTE structure [16]

In accordance with the result of IQ sampling, the phase of the CTE at the sampling instant may be determined by formula (1).

$$\varphi' = \operatorname{atan}\left(\frac{\varrho}{\iota}\right),\tag{1}$$

where  $\varphi'$  is the phase of antenna, *I* is the value of inphase component, and *Q* is the value of quadrature component.

The angle is measured by formula (2).

$$\theta = \arccos(\frac{\psi\lambda}{2\pi d}) \tag{2}$$

Where  $\theta$  is angle of arrival,  $\psi$  is the phase difference between the antennas, *d* is the distance between the antennas, and  $\lambda$  is the wavelength of the Bluetooth signal.

Before measuring the tag location, it is also necessary to decompose the AoA into elevation and azimuth angles. Figure 5 shows an example of measuring the tag location by elevation and azimuth angles.



$$x = h \times \tan\left(\frac{e \times \pi}{180}\right) \times \cos\left(\frac{a \times \pi}{180}\right), \qquad (3)$$

$$y = h \times \tan\left(\frac{e \times \pi}{180}\right) \times \sin\left(\frac{a \times \pi}{180}\right), \tag{4}$$

where x and y are the coordinate of the tag location, h is the height between anchor and the tag, e is the elevation angle, and a is the azimuth angle.

#### 4. ANOMALY DETECTION METHODS

AD methods are processes that find the outliers of a dataset. These outliers must be eliminated before analysis. Because the outlier make errors in the result of locating. In order to reduce these errors, eliminating outliers at the earlier stage may contribute to accurate indoor positioning.

Figure 6 shows errors expressed over time at a certain location. the error. The errors refer to the Euclidean distance between the estimated user location and the actual user location. Figure 6 shows errors are unstable.



Figure 6: Errors of a certain user location [12]

In order to reduce these errors, outliers need to be removed. GESD, LevelShift, and Persist methods are used to eliminate outliers. Figure 7 briefly presents the procedure of improving the accuracy of AoA.

ISSN: 1992-8645 E-ISSN: 1817-3195 www.jatit.org 35 (a) Measure the actual user location (b) Obtain I/Q data of each user location from the AoA anchor (c) The AoA data are transferred by I/Q data (d) Apply each AD method to the AoA data (e) Estimate the user locations using the normal AoA data



The procedure of improving the accuracy of AoA is introduced. Step 1, the actual user location is measured in (a). Step 2, the AoA anchor receives Bluetooth signal from tag at each user location. I/Q data is measured by AoA anchor in (b). Step 3, I/Q data is used to get the AoA data in (c). Step 4, the AD is applied to the AoA data which contains outliers in (d). Step 5, the user location is estimated by the normal AoA data. In the following sections, three AD methods are introduced.

#### 4.1 GESD method

GESD is only used when the data set are following the normal distribution. GESD is a simple statistical approach in a univariate data set. Figure 8 shows the different normal distribution in the various degrees of freedom (df). The parameter alpha indicates the distribution can be changed manually in GESD. Alpha can make the shape of normal distribution become different. For example, Figure 9 depicts the data set which indicates example of applying GESD when the parameter alpha is 0.3. The red dots indicate anomalies in the graph.



Figure 8: Distributions with different shapes depending on the degree of freedom [13]

#### 4.2 LevelShift method

The following context introduces LevelShift method. InterQuantileRange and Threshold method are firstly introduced because LevelShift method is made of these two methods.



The box plot is a statistical approach to eliminate outliers. In a certain data, the most special values are the median, maximum, and minimum. The boxplot defines the value corresponding to 0% as the minimum, 25% as  $Q_1$ , 50% as the median, 75% as Q<sub>2</sub>, and 100% as the maximum. This example is easily illustrated by the Figure 10.



Figure 10: Boxplot for explaining InterQuantileRange method [17]

IQR (InterQuantileRange) is a range of  $Q_3$  and  $Q_1$ . Upper Value and Lower Value are determined by IQR,  $Q_1$ ,  $Q_3$ ,  $C_1$ , and  $C_2$ . The parameter of  $C_1$  and  $C_2$ values can be set manually. If a data is not within the range of Upper Value and Lower Value, it is detected as anomalies.

$$IQR = Q_3 - Q_1, (5)$$

Upper Value =  $IQR + Q_3 \times C_1$ , (6)

$$Lower \ Value = IQR - Q_1 \times C_2 \tag{7}$$

Threshold method is a fundamental method to detect anomalies by setting the maximum and the minimum value.

LevelShift is the method by merging IQR method and Threshold method. There are parameter c and window that can be set manually. Parameter c sets a range of normal values. The parameter *window* is the size of the time to compare the data. The result of AD is changed by parameter *c* and *window*.

For using parameters optimally, they can be set manually. An optimal parameter can be found as follows: Parameter c can be tried 0 and 10 first. The <u>31<sup>st</sup> October 2022. Vol.100. No 20</u> © 2022 Little Lion Scientific



E-ISSN: 1817-3195

www.jatit.org result of using LevelShiftAD with c of 10 is better than 0. Next step is trying another c like 6. Then it can be found 6 is optimal value of c. The procedure of determining window is like to parameter c. The result of using LevelShift method with parameter c of 6.0 and window of 5 is shown in Figure 11.



#### 4.3 Persist Method

ISSN: 1992-8645

Persist method also uses Threshold method and IQR method like LevelShift method. Persist Method uses past time set to detect anomalies. Persist method only check one previous value, which is good at capturing additive anomaly in short-term scale, but not in long-term scale because it is too near-sighted [8]. Figure 12 shows an example of Persist method with c is 3.0. The procedure of determining c is like that of GESD.



3.0 [11]

#### **INDOOR** 5. AOA POSITIONING **EXPERIMENT**

Figure 13 shows a user location in the map. The red dot represents the user location, and the black dot represents the location of AoA anchor which is in the center of classroom.

Figure 14 shows the experimental environment of this paper. A classroom of 7.6 m in length and 7.85 m in width is chosen for testing. AoA anchor is attached on the ceiling in the classroom, 2.52 m above the floor. A person takes tag standing in a user

location. This paper uses a USB cable to obtain AoA data from the AoA anchor. The AoA demo python program transfers the AoA data to the user location.



*Figure 13:* A user location with x=0, y=2.7 in the тар



*Figure 14: Experiment environment [12]* 

Figure 15 shows the AoA anchor used in this experiment. The AoA anchor has 12 antennas. The Bluetooth signal from the tag can be received by the AoA anchor from all directions. The location of the tag can be measured by the AoA data from the AoA AoA anchor. The anchor uses Nordic semiconductor's nRF52833 chip.



Figure 15: AoA anchor [12]

Figure 16 shows the data from the AoA anchor. ME is the unfiltered elevation angle data. MA is the unfiltered azimuth angle data. In this paper, the function of AD methods which is similar with filter are used for detecting and eliminating the outliers in the AoA data.

<u>31<sup>st</sup> October 2022. Vol.100. No 20</u> © 2022 Little Lion Scientific

ISSN: 1992-8645	www.jatit.org	E-ISSN: 1817-3195
Figure 17 shows ME and MA ang	gles in a certain Most of MA angles	are determined as normal value
	1 1	1 0 · · T · · · 11

location. MA angles are fluctuated severely between 0 and 360 degrees. Even ME angles are also fluctuated between 0 and 90 degrees. Anomalies must be detected and eliminated by the AD method. In this paper, several AD methods are applied to ME and MA to detect outliers. GESD, IQR, and Persist AD are tested in this paper.

#### 5.1 The test result of GESD

GESD can set the parameter alpha set from 0 to 1 manually. In GESD, parameter alpha effects the result of GESD. In this test result of GESD, however, parameter alpha doesn't affect the result. Because the number of the outliers which are determined by GESD is smaller than the number of actual outliers. If the pattern of outliers is appeared in ME and MA, GESD is recognized the outlier as normal value. Figure 18 shows ME angles in time order at a user location. An outlier between 10 and 20 degrees, which appears occasionally, is well found by GESD. While an ME angle is 89 degrees, which appears continuously, is not recognized as anomalies by GESD. This is a disadvantage of GESD.



Figure 16: Data collected by the AoA demo program [12]

Even though, GESD has a disadvantage, in some cases it can detect outliers very well like Figure 19. Figure 19 shows the result of GESD in MA angles. because there are little fluctuations. In this case, all of parameters in GESD shows the same result as Figure 18, and Figure 19.

Figure 20 and Figure 21 show the result of using GESD in ME and MA angles without outliers. The minimum value of y-axis is increased after the outlier is eliminated. It should be reflected in the indoor positioning by maximizing this advantage. Because GESD reduces the influence of outliers on the ME and MA angles.

Figure 17: In a certain location angle data are



collected by ME, MA format [12]

#### 5.2 The test result of LevelShift

LevelShift is a method to find out whether there has been a change by comparing specific time point value and the value before that. Figure 22 shows MA angles in time order at a user location. There are parameters c and window which are determined manually until they are set as optimized. In Figure 22 it looks like detecting anomalies better than GESD, but it has a problem. If the present value is regarded as an anomaly when compared with the past value, LevelShift determines both the present value and the past value as anomalies. Therefore, even if one of the two angles is normal, one wrongly determined. The larger the window parameter, the more normal angles are determined as anomalies. The data fluctuate greatly as it is shown in Figure 22, and Figure 23. Figure 22 and Figure 23 show the problems caused by non-optimized parameters. Figure 22 shows that LevelShift is applied in MA with c is 0.001 and window is 1. Figure 23 shows that LevelShift is applied in MA with c is 0.8 and window is 5. As it is shown in both Figure 22 and Figure 23, the smaller the parameter c, the smaller the change is found as an anomaly. If any one of them is the abnormal value, parts of MA angles are determined by anomalies.



Figure 18: The example of using GESD in ME at a user location



Figure 19: The example of using GESD in MA at a user location



Figure 20: The example of using GESD in ME out of outliers at a user location



Figure 21: The example of using GESD in MA out of outliers at a user location





*Figure 22: The example of using LevelShift in MA with parameter c is 0.001 and window is 1 at a user location* 



Figure 23: The example of using LevelShift in MA with parameter c is 0.8 and window is 5 at a user location

Figure 24 shows a graph without outliers at a user location and the same parameters as Figure 22. Similarly, Figure 25 shows a graph without outliers at a user location and the same parameters as Figure 23. The value of 359 degree is appeared frequently in MA. So LevelShift determines 359 degree as the normal value.



Figure 24: The example of using LevelShift in MA with parameter c is 0.001 and window is 1 out of the outliers at a user location



Figure 25: The example of using LevelShift in MA with parameter c is 0.8 and window is 5 out of the outliers at a user location

© 2022 Little Lion Scientific

## ISSN: 1992-8645

www.jatit.org

#### 5.3 The test result of Persist

Persist has a parameter *c* and a parameter *window*, but in contrast to LevelShift, the parameter window is fixed to 1. By only changing parameter c, Persist characteristics can be known. Like LevelShift it compares with past data to determine whether a certain data is normal or not. Figure 26 shows that Persist is applied in MA with c is 0.1. Figure 27 shows that Persist is applied in MA with c is 0.8. In the Figure 26 and Figure 27, it seems that the smaller the c value, the more anomalies are detected by LevelShift. In fact, the most optimized parameter could not be found because normal angles are recognized as anomalies in Figure 26 and Figure 27.



Figure 26: The example of using Persist in MA with parameter c is 0.1 at a user location



Figure 27: The example of using Persist in MA with parameter c is 0.8 at a user location

Figure 28 shows a graph without outliers at a user location and the same parameters as Figure 26. Similarly, Figure 29 shows a graph without outliers at a user location and same parameters as Figure 27. Like GESD, LevelShift, and Persist detect same

anomalies of appearing in MA as normal value. For example, the value of 359 degree appears in MA so many times. Three methods detect 359 degree as normal value.



Figure 28: Persist is applied in MA with parameter c is 0.1 out of the outliers at a user location

ISSN: 1992-8645 E-ISSN: 1817-3195

Figure 29: Persist is applied in MA with parameter c is 0.8 out of the outliers at a user location

Figure 30 shows the RMSE after applying each ADs and removing the outliers. The RMSE of GESD is 13.5 m. The RMSE of LevelShift is 12.2 m. The RMSE of Persist is 14.2 m. The RMSE of raw data without filtering is 13.9 m. LevelShift, the best performance, shows the improvement by 14% in the RMSE. GESD also shows the improvement by 5%.

In the case of Persist, however, it gets worse by 2%. Each ADs determine that some outliers such as 359 degrees in ME angle, and 89 degrees in MA angle as normal because it appears 11 times. These anomalies are not filtered and eventually included in the estimated user location. That causes error estimating the RMSE values in every AD method.



Figure 30: RMSE In A Certain User Location Comparing With Applying Each Ads And Raw Data At A User Location

#### 6. CONCLUSION

This study focuses on improving the accuracy of AoA indoor positioning technology based on Bluetooth 5.1. The detection of outliers is a key method of improving the accuracy of AoA indoor positioning technology. GESD, LevelShift, and Persist are tested to detect outliers in this paper. The result shows that GESD has 5% improvements, and Levelshift shows 14% better than raw data in the RMSE. But Persist shows 2% worse.

There are still outliers still not filtered. The ship's environment may occur lots of noises including outliers appear in the AoA data. Therefore, it is necessary to research and develop another AD that can detect outliers as much as possible. That is more suitable for AoA data in the data preprocessing stage for the ship's environments.

#### ACKNOWLEDGEMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant number: 2021R111A3056125).



© 2022 Little	Lion Scientific
ISSN: 1992-8645	.jatit.org E-ISSN: 1817-3195
REFERENCES	Wireless Communications and Networking
	Conference (WCNC). IEEE, 2021.
[1] Monfared S, Nguyen T H, Petrillo L, et al. Expe	[11] ARDUNO, Detector — ADTK 0.6.2
rimental demonstration of BLE transmitter posi	documentation
tioning based on AOA estimation[C]//2018 IEE	[12] Hyeongseol Shin, Qianfeng Lin, Jooyoung Son.
E 29th Annual International Symposium on Per	An analysis of the accuracy of Bluetooth 5.1-
sonal, Indoor and Mobile Radio Communicatio	based AoA indoor positioning algorithms for
ns (PIMRC). IEEE, 2018: 856-859.	ship passengers'
[2] Cominelli M, Patras P, Gringoli F. Dead on arriv	health[C]/KoreaMultimediaSociety, 2022, 1-4.
al: An empirical study of the Bluetooth 5.1 posi	[13] oneofakindscene. SH-ESD OUTLIER
tioning system[C]//Proceedings of the 13th inte	DETECTION. https://walog.jo/@oneofakindscene/SH
rnational workshop on wireless network testbed	$ESD \ \%EC\% \ 0D\% \ D\% \ C\% \ C\% \ S3\% \ S1\% \ EC\% \ D\% \ D\% \ C\% \ S1\% \ S1\%$
s, experimental evaluation & characterization. 2	%98-%ED%83%90%EC%47%80 07-23-
019: 13-20. [2] Ey V. Zhang H. Zhang H. Divataath A a A. Dasitia	2021
ning Based on Backscatter Technology[C1//202	[14] H. Liu, H. Darabi, P. Baneriee and J. Liu,
0 IEEE 20th International Conference on Com	"Survey of Wireless Indoor Positioning
munication Technology (ICCT) IEEE 2020: 5	Techniques and Systems," in IEEE Transactions
59-565.	on Systems, Man, and Cybernetics, Part C
[4] Toasa F A, Tello-Oquendo L, Peńafiel-Ojeda C	(Applications and Reviews), vol. 37, no. 6, pp.
R, et al. Experimental demonstration for indoor	1067-1080, Nov. 2007, doi:
localization based on AoA of Bluetooth 5.1 usi	10.1109/TSMCC.2007.905750.,
ng software defined radio[C]//2021 IEEE 18th	[15] Vigneshdm1990. IQ phasor diagram.
Annual Consumer Communications & Network	https://en.wikipedia.org/wiki/In-p
ing Conference (CCNC). IEEE, 2021: 1-4.	hase_and_quadrature_components#/media/File
[5] Paulino N, Pessoa L M, Branquinho A, et al. Eva	:IQ_phasor_diagram.svg. January 26, 2020
luating a Novel Bluetooth 5.1 AoA Approach fo	[16] Nordic Semiconductor, Direction Finding nwP-
r Low-Cost Indoor Vehicle Tracking via Simula	[17] Krishnan Srinath Anamoly Detection:
tion[C]//2021 Joint European Conference on N	Techniques to detect outliers
etworks and Communications & 6G Summit (E	https://towardsdatascience.com/anamoly-
[6] Haji Akhandi Mayhadi 7, Salimihani M. Maham	detection-techniques-to-detect-outliers-
madi A et al Bluetooth low energy and CNN-h	fea92047a222 June 22, 2020
ased angle of arrival localization in presence of	,
Rayleigh fading[C]//ICASSP 2021-2021 IEEE	
International Conference on Acoustics, Speech	
and Signal Processing (ICASSP). IEEE, 2021:	
7913-7917.	
[7] Hajiakhondi-Meybodi Z, Salimibeni M,	
Mohammadi A, et al. Bluetooth low energy-	
based angle of arrival estimation in presence of	
Rayleigh fading[C]//2020 IEEE International	
Conference on Systems, Man, and Cybernetics	
(SMC). IEEE, 2020: 3395-3400.	
[8] Hong, Ke, et al. "A Learning-Based AoA	
Estimation Method for Device-Free	
(2022)	
[9] Wang Zhonggin et al "Accurate AoA	
Estimation for RFID Tag Array With Mutual	
Coupling." IEEE Internet of Things Journal 9 15	
(2022): 12954-12972.	
[10] Smaoui, Nour, Milad Heydariaan. and	
Omprakash Gnawail. "Single-antenna aoa	
estimation with uwb radios." 2021 IEEE	