

# Implementation K-Nearest Neighbor Algorithm in Searching Location Books in Library Statically Based on RFID

Ahmad Fali OKLILAS<sup>1\*</sup>, Fajri Aulia RACHMAT<sup>2</sup>, and Purwita SARI<sup>3</sup>

<sup>1,2</sup>*Department of Computer Engineering Faculty of Computer Science Universitas Sriwijaya, Palembang, Indonesia*

<sup>3</sup>*Study Program of Informatics Management Faculty of Computer Science Universitas Sriwijaya, Palembang, Indonesia*

*\*Corresponding author: fali@ilkom.unsri.ac.id*

## ABSTRACT

Radio Frequency Identification (RFID) is a technology for determining an object using electromagnetic waves (radio waves) through a device called a tag. This identification is carried out using a reader, tag, and antenna. To maximize the performance of RFID technology in a room with a room scale that is not too large, the K Nearest Neighbor (KNN) algorithm is used to calculate the error value generated from the reader and antenna that measures tags and is applied in an automatic counting program designed using the Python programming language. After testing and experimenting on target tags using each of the 4 reference tags, the test and experiment results do not differ greatly from the actual coordinates of the detected tags for target tag A with a percentage error coordinate value of ( $x = 6.3\%$ ), ( $y = 3\%$ ), and ( $z = 0.01\%$ ). The target Tag A error values of 4 reference tags are ( $x = 6.3$  cm), ( $y = 1.05$  cm), and ( $z = 0.4$ cm). Meanwhile, the coordinate error values ( $x_e, y_e, z_e$ ) for target B Tag with the percentage error coordinate values of ( $x = 3\%$ ), ( $y = 4\%$ ), and ( $z = 0\%$ ). The target Tag B error values of 4 reference tags are ( $x = 3.3$  cm), ( $y = 2.8$ cm), and ( $z = 0$  cm).

**Keywords:** *identification, K-Nearest Neighbor (KNN) algorithm, RFID*

## Introduction

Radio Frequency Identification (RFID) is a technology to identify an object using electromagnetic waves through a device called a *tag*. Compared to identification technology *barcode* traditional, RFID is superior and efficient because it does not require assistance to access data, has a greater reading distance with the *reader* able to communicate with multiple *tags* [1]. RFID labels contain information that is stored electronically and can be read up to several meters away. RFID reader systems do not require direct contact such as barcode systems [2].

Problems in the search system for the location of books in the library during this time manually search directly on the bookshelves, this is what causes less efficient ways to find the location of books in the library. For that we need a way to find the location of books with a system that is faster and easier. The method used is based on previous literature by placing RFID tags in the books sought. Then the RFID reader is used to find the location of the RFID tag attached to the book. By using this RFID system, it will be easier to find the location of books in the library. This search is still static, not dynamic, so the book that is sought in this case study is a book that is being static is in a location, not the location of a book that is moving. In this paper the locations that will be searched ordinarily are X, Y, and Z. So that it will be known where the location

of the book is in a three-dimensional space. The results will be sought also how much the value of the error or the percentage error. This error value aims to test how accurate the system has been made. The algorithm used in the calculation is the K-Nearest Neighbor (KNN) algorithm.

## Literature Study

Study Literature and studies are stages of finding, reading, studying and understanding references related to this research as support for writing reports. The purpose of this literature study is to obtain sufficient references to the work process and methods for completing research. The literature used in this study comes in part from several scientific journals, and reference books.

## Radio Frequency Identification (RFID)

Radio frequency identification (RFID) is a technology that uses communication via electromagnetic waves to change data between terminals and objects such as products of goods, animals, or humans for the purpose of identification and trace tracking through the use of a device called RFID tag [3].

There are 3 main components of RFID, namely tags, readers, and antennas. The RFID tags are divided into 3

types namely; Passive Tag; Semi-Passive Tag; Active Tag. Passive tags do not have batteries or sensors, waves sent by the antenna reader make a passive tag system work [4]. Passive tags are that use electromagnetic energy received from transmissions reader to reply to readers. On the other hand, tags passive are usually cheaper, smaller, and lighter than other types of tags, with some attractive advantages for many RFID applications [5].

Antennas on RFID function to transmit radio frequency signals between the reader and the tag. Data or information is then encoded into radio frequency waves through the transmission channel. So, the antenna has a very important role in the RFID system [6].

### Received Signal Strength Indicator (RSSI)

Received Signal Strength Indicator (RSSI) is a standard feature in most localization solutions and is defined as a voltage signal strength indicator pin that is received on a radio signal. However, RSSI is considered a key parameter for estimating object coordinates and as such, it is very important for accurate localization. Radio frequencies commonly used in RFID include 120-150 kHz in *Low Frequency* (LF), 13.56 MHz, *High Frequency* (HF), 433MHz, 868-870 MHz, and 902-928 MHz in *Ultra High Frequency* (UHF), and 2.4-5.8 GHz *Microwave frequency* (MW) [7].

### K-Nearest Neighbor Algorithm (KNN)

The *K-Nearest Neighbor* (KNN) algorithm is a method for classifying objects based on learning data that is the closest distance to the object. The *K-Nearest* who are supervised where the results of a query instance new classified by the majority category of KNN [8]. The *K-Nearest Neighbor* or KNN, the main parameters used are the RSSI of the target tag that wants to know its estimated position (RSSI<sub>target</sub>), the RSSI reference tag that has its position (R<sub>i</sub>) and the position of a reference tag (x<sub>i</sub>, y<sub>i</sub>, z<sub>i</sub>) and will get the approximate position of the target tag (x<sub>e</sub>, y<sub>e</sub>, z<sub>e</sub>). Some equations are used in the *K-Nearest Neighbor* algorithm or KNN in determining the approximate position [9].

Below is the equation of the Euclidean Distance Value:

$$E_i = \sqrt{\sum_{i=1}^n (t_g - t_r)^2}$$

(1)

E<sub>i</sub> is the value of a Euclidean, t<sub>g</sub> is the RSSI value of two antennas on the tag targets A and B, t<sub>r</sub> is the RSSI value of two antennas on the Tag reference for each tag targets A and B and E<sub>i</sub> depending on the number of reference tags used in this study.

$$W_i = \frac{\frac{1}{E_i^2}}{\sum_{i=1}^k \frac{1}{E_i^2}}$$

(2)

W<sub>i</sub> is a weighting factor for the number i neighbors (neighbors) Reference tags, W<sub>i</sub> depends on the value of E of each target Tag in the test environment where W<sub>i</sub> is a function value E of neighbor (neighbor) nearby. After the value of W<sub>j</sub> obtained, then find the coordinates estate of Tag targets of the coordinate value Tag reference:

$$(x, y, z) = \sum_{i=1}^k W_i (x_i, y_i, z_i)$$

(3)

From equation (3) above, the coordinate value and the number of K (error) what is the smallest error. The value of W (weight) of each tag previously obtained using equation (3), while x, y and z are coordinate estimations of the reference tag. The Algorithm process is *K-Nearest Neighbor* complete and the position is likely to be obtained then the target tag position error can be calculated based on the tracking and localization system using this algorithm *K-Nearest Neighbor* with the current tag position using equation (4) discussed earlier as shown below:

$$error = \sqrt{(x - x_e)^2 + (y - y_e)^2 + (z - z_e)^2}$$

(4)

### Percentage of Error and Current Distance

In the last process, the percentage of error will be calculated using the smallest K value in each experiment on each target tag. This error percentage value will be calculated using 2 methods, the first is to find the middle value of the value (equation (5)), then use the equation (6) as shown below:

$$middle\ value = \frac{experiment\ 1\ coordinate + experiment\ 2\ coordinate + \dots}{number\ of\ experiment\ (n)}$$

(5)

$$\%error = ((middle\ value - actual\ value) / middle\ value * 100\%)$$

(6)

To calculate the current distance from the error percentage the value of the coordinate value of the axis (x<sub>e</sub>, y<sub>e</sub>, z<sub>e</sub>), the percentage value obtained is calculated with the current value as calculated in the equation (7) below:

$$actual\ error\ distance = (\%error * x\ axis)(\%error * y\ axis)(\%error * z\ axis)$$

(7)

**METHODOLOGY**

The methodology used to complete this final project goes through several stages that follow the framework that was previously designed. The stages of the methodology include the stages of carrying out a literature study or looking for literature related to research, stages in determining the workspace and preparing the software used in the RFID system reader, the initial data retrieval stage, the calculation phase using the K-Nearest Neighbor Algorithm and determining the best position. Then the results will be analyzed to draw conclusions.

**RESEARCH FRAMEWORK**

The framework includes determining the need for a Tag tracking system in an RFID system. The form of the framework in this research is to conduct Literature and Literature Study, then by determining the parameters of the place of research. Next, determine the hardware requirements namely Impinj SpeedWay R240 Reader, Impinj Antenna, and UHF Tag. Next, perform initial data retrieval by retrieving RSSI values using the Impinj SpeedWay R240 reader manually, and finally, performing data processing using the KNN Algorithm, looking for error values, and analyzing the results.

**WORK ENVIRONMENT INITIALIZATION AND TAG PLACEMENT**

Stages will explain the work environment and the placement of the target Tag and reference tags that will be used in this study. The working environment used is room (2mx1.5mx1,2m) in which there is a table with a length of

70cm, width 30cm and height 40cm, and a chair with a length of 20cm, width 20cm and width 20cm which will be used as a place to put a target tag and reference tag, which will be symbolized by coordinates (x, y, z) x for the length of the room, y for width, and z for height of the room. Research will use two antennas Reader, 8 Tag 2 Tag references and targets. Reference tags will be placed in the table, and a chair fields, their positions are in each corner of the table and chair fields. Reference tags that are used in the table are 4 pieces, and in the seat field also 4 pieces. The following is an illustration of the reference tag placement in Figure 1.

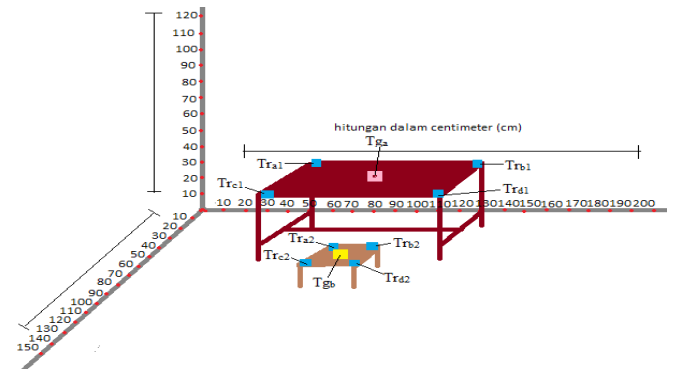


Figure 1. Illustration placement of the target tag and a reference tag

**RESULTS AND DISCUSSION**

In this experiment and research, 6 tests were conducted on each target tag, tables 1 and 2 below are RSSI value testing results that have been done automatically calculate the smallest error value for the reference tag and target tag:

Table 1. RSSI value Target A regeneration tag in dBm

	RSSI value from : Reference Tags (Tr1) on target A Tag	Antenna 1 (dBm)	RSSI value from : Reference Tags (Tr1) on target A Tag	Antenna 2 (dBm)
a	Tr <sub>a1A</sub>	-53	Tr <sub>a1B</sub>	-49
b	Tr <sub>b1A</sub>	-62	Tr <sub>b1B</sub>	-63
c	Tr <sub>c1A</sub>	-58	Tr <sub>c1B</sub>	-55
d	Tr <sub>d1A</sub>	-49	Tr <sub>d1B</sub>	-54

Table 2. RSSI value for target A Tag in dBm

	RSSI value from: Target A Tag (Tg)	Antenna 1 (dBm)	RSSI value from: target A Tag (Tg)	Antenna 2 (dBm)
a	Tg <sub>a1</sub>	-55	Tg <sub>a2</sub>	-54

The following table 3 and 4 are the results of testing the RSSI value after an automatic calculation obtains the

smallest error value for the reference Tag and target Tag B:

Table 3. RSSI value Reference tag for Target B in dBm

RSSI value from : Reference Tags (Tr2) on target B Tag	Antenna 1 (dBm)	RSSI value from : Reference Tags (Tr2) on target B Tag	Antenna 2 (dBm)
Tr <sub>a2A</sub>	-52	Tr <sub>a2B</sub>	-59
Tr <sub>b2A</sub>	-56	Tr <sub>b2B</sub>	-62
Tr <sub>c2A</sub>	-58	Tr <sub>c2B</sub>	-65
Tr <sub>d2A</sub>	-60	Tr <sub>d2B</sub>	-63

Table 4. RSSI value for target B Tag in dBm

RSSI value from: target B Tag (Tg)	Antenna 1 (dBm)	RSSI value from: target B Tag (Tg)	Antenna 2 (dBm)
<b>b</b> Tg <sub>b1</sub>	-52	Tg <sub>b2</sub>	-50

### Calculation of Euclidean value € an calculation Weight Factor value (W)

To calculate this Euclidean value used equation (1), which has obtained the results of the Euclidean value obtained from the RSSI value and for the calculation of the Weight value (W) use equation (2) for the target Tag A is in table 5 below this:

Table 5. Euclidian Value (E) and Weight Factor Value (W) for the target tag A

Next , to calculate the weight (W) value for target B tags is in table 6 below :

Table 6. Euclidian Value (E) and Weight Factor Value (W) For the target Tag B

	$E_i$	$\frac{1}{E_i^2}$	$W$ (Weight)
<b>a</b>	9	0,02	0,75
<b>b</b>	13,41	0,01	2,66 <sup>10-2</sup>
<b>c</b>	19	0,01	0,11
<b>d</b>	15,26	7,67 <sup>10-5</sup>	0,13
$\Sigma$		0,03	

## Coordinates and errors

Find coordinates and errors using equations (3) and (4). The following table 7 is the coordinate value estimates and the error value in Tag research targets:

Table 7. Estimation Position and error for each K (1 to 4) Target A tag

<b>k</b>	<b><math>x_e</math></b>	<b><math>y_e</math></b>	<b><math>z_e</math></b>	<b>error (cm)</b>
1	31,55	9,02	18,03	76,44
2	45,63	11,03	22,05	62,07
3	52,31	15,8	25,87	53,32
4	101,77	33,46	40	2,34

Next, table 8 is the estimated coordinate value and error value in the target B Tag in the research :

Table 8. Estimation Position and error for each K (1 to 4) Target B tag

<b>k</b>	<b><math>x_e</math></b>	<b><math>y_e</math></b>	<b><math>z_e</math></b>	<b>error (cm)</b>
1	51,47	30,88	10,29	71,06
2	80,33	50,12	15,10	36,05
3	104,56	64,66	19,95	7,637
4	104,86	64,86	20	7,27

## K-Nearest Neighbor (KNN) dan errors

For the error value obtained from the reference Tag for the target A tag with the coordinate position ( $x_e, y_e, z_e$ ) = (100 ; 35 ; 40) are for experiment 1 (96.16 ; 31.03 ; 40.18), experiment 2 (87.68 ; 24.26 ; 40), experiment 3 (101.77 ; 33.46 ; 40), experiment 4 (92.63 ; 38.27 ; 40), experiment 5 (92.63 ; 38.27 ; 40), experiment 6 (93.24 ; 38.03 ; 40). From the error value, the calculation of the percentage value of the error value will be done again. The target tag obtained from the value  $K = 4$  in each experiment. The middle value of experiment 1 through experiment 6 as in equation (5), will be calculated by equation (6) for the percentage value of coordinate error with respect to the axis ( $x_e, y_e, z_e$ ) on the target A Tag is (94.02) (33.9) (40,03). To find the percentage error value in the coordinate value against the axis ( $x_e, y_e, z_e$ ) for the actual value used equation (6) with results (6.3%) (3%) (0.01%). To calculate the actual distance of the percentage error value in the coordinate value against the axis ( $x_e, y_e, z_e$ ) on the target A Tag, the percentage value obtained is calculated with the actual value in units of centimeters such as the calculation of equations (7) with result (6.cm) (1.05cm) (0.4cm).

For the target A tag with the coordinate position ( $x_e, y_e, z_e$ ) = (110; 70; 20) for experiment 1 (116.82; 76.3; 20.01), experiment 2 (105.01; 65.01; 20), experiment 3 (106.25; 66.25; 20), experiment 4 (106.25; 66,25; 20), experiment 5 (102.73; 62,73; 20), experiment 6 (104.86; 64.86; 20). From the error value, calculation of the percentage value of the error value will be done again. The target tag obtained from the value  $K = 4$  in each experiment. The middle value of experiment 1 through experiment 6 as in equation (5), will be calculated by equation (6) for the percentage value of coordinate error with respect to the axis ( $x_e, y_e, z_e$ ) on the target B Tag is (106.99) (66.9) (20 for the percentage value of coordinate error with respect to the axis ( $x_e, y_e, z_e$ ) for the actual value used equation (6) with results (3%) (4%) (0 %). While, To calculate the actual distance of the percentage error value in the coordinate value against the axis ( $x_e, y_e, z_e$ ) on the target A Tag, the percentage value obtained is calculated with the actual value in units of centimeters such as the calculation of equations (7) with result (3.3.cm) (2.8cm) (0cm).

## SUMMARY

From the estimated position of coordinates (xe, ye, ze) obtained from the target Tag A and the target B tag are for the target Tag A with the largest value with the value of the Nearest Neighbor (K = 1), with coordinate values (21.65; 6.19; 12.37). And for the smallest value with the value of the Nearest Neighbor K (K = 4) with coordinate values (101.77; 33.47; 40). For K with sequence 4 there is the closest to the current coordinates of the target Tag A, that is (100; 35; 40). As for the target B Tag with the largest value with the value of Nearest Neighbor K (K = 1), with coordinate values (51.47; 30.88; 10.29). For the smallest value with the value of Nearest Neighbor K (K = 4), with coordinate values (104.86; 64.86; 20). This research has succeeded in finding the position of the target tags A and B by using the KNN algorithm, from the calculation of the error for the target tag A, obtained the value of the coordinate error (xe, ye, ze) for the target A tag with a percentage error coordinate value of (x = 6.3%), (y = 3%) and (z = 0.01%) Tag target A error values of 4 reference tags are (x = 6.3 cm), (y = 1.05 cm), and (z = 0.4cm). Meanwhile, the coordinate error values (xe, ye, ze) for target B Tag with the percentage error coordinate values of (x = 3%), (y = 4%), and (z = 0%). The target Tag B error values of 4 reference tags are (x = 3.3 cm), (y = 2.8cm), and (z = 0 cm).

## REFERENCES

- [1] A. F. On and I, 2008, "Radio-Frequency Identification ( RFID )," *Radio-Frequency Identif. ( RFID )*, no. **138**,
- [2] T. Karygiannis, B. Eydt, G. Barber, L. Bunn, and T. Phillips, 2007, "Guidelines for securing radio frequency identification (RFID) systems- Recommendations of the," *Natl. Inst. Stand. ....*,
- [3] F. Communication, 2012, "RFID Handbook : Fundamentals and Applications in Contactless Smart Cards , Radio ... Page 1 of 2 RFID Handbook : Fundamentals and Applications in Contactless Smart Cards , Radio Frequency Identification and Near RFID Handbook : Fundamentals and Application," pp. **2–3**.
- [4] L. X. Zheng, Z. Xie, and X. Cen, 2013, "A compact RFID reader antenna for UHF near-field and far-field operations," *Int. J. Antennas Propag.*, vol. **2013**.
- [5] A. M. Shaari and N. S. M. Nor, 2017, "Position and Orientation Detection of Stored Object Using RFID Tags," *Procedia Eng.*, vol. 184, pp. **708–715**.
- [6] R. Patel, N. Patel, and M. Gajjar, 2012, "Online Students ' Attendance Monitoring System in Classroom Using Radio Frequency Identification Technology : A Proposed System Framework," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 2, no. 2, pp. **61–66**.
- [7] G. Çaliş, B. Becerik-Gerber, A. B. Göktepe, S. Li, and N. Li, 2013, "Analysis of the variability of RSSI values for active RFID-based indoor applications," *Turkish J. Eng. Environ. Sci.*, vol. 37, no. 2, pp. **186–210**.
- [8] E. Suranta B. S, 2014, "Algoritma K-Nearest Neighbor (KNN)," *Igarss 2014*, vol. 72, no. 1, pp. **1–5**.
- [9] W. Yustanti, 2012, "Algoritma K-Nearest Neighbour untuk Memprediksi Harga Jual Tanah," *J. Mat. Stat. dan komputasi*, vol. 9, no. 1, pp. **57–68**.