# AASEC 2019

aasec

## AASEC 2019 : Please Activate Your Account

Eksternal  Kotak Masuk ×

**AASEC 2019** aasec2019.automail@upiconf.org lewat srv148.main-hosting.eu
kepada saya, aasec

Inggris > Indonesia  Terjemahkan pesan

Dear Dr. Mujiarto Mujiarto,

Your registration to AASEC 2019 has been completed,
please open (copy-paste to browser) the following link to activate your account:

http://aasec2019.upiconf.org/kfz/pages/activate.php?q=a9b151c416a4ce83e7ae9eae059839fd

Thank you.
Best regards,

AASEC 2019 Organizing Committee
Website : http://aasec.conference.upi.edu/2019
Email : aasec@upi.edu

----------------------------------------------------------------
http://konfrenzi.com | Web Systems for Scientific Conferences

## AASEC 2019 : Your Registration has been Approved

Eksternal  Kotak Masuk ×

**AASEC 2019** aasec2019.automail@upiconf.org lewat srv148.main-hosting.eu
kepada saya, aasec

Inggris > Indonesia  Terjemahkan pesan

Dear Dr. Mujiarto Mujiarto,

Your Registration has been Approved.
User ID: USER-886
Please use this "User ID" in all correspondence (instead of your name).

Email: mujiarto@umtas.ac.id
Login Code: vU7oFXImle

You need the "Login Code" to login to our site, so please do not delete this email.

Please login to submit your abstract and paper.

Thank you.
Best regards,

AASEC 2019 Organizing Committee
Website : http://aasec.conference.upi.edu/2019
Email : aasec@upi.edu

----------------------------------------------------------------

aasec

50 dari banyak

## Notification of AASEC Abstract Acceptance

Eksternal | Kotak Masuk ×

**Annual Applied Science Engineering Conference** <aasec@upi.edu>
kepada bcc: saya

Rab, 20 Mar 2019 23.31

Inggris > Indonesia | Terjemahkan pesan | Nonaktifkan untuk: Inggris ×

Dear Participants,

We would like to thank you for your interest and enthusiasm in the 4th AASEC 2019 that will be organized this year by Universitas Pendidikan Indonesia, Universitas Warmadewa and also the other co-hosts.

A total of 1.000 participants have registered to this year's conference, including you. This leads an objective selection of the abstracts to be presented on the conference.

In this occasion, we would like to congratulate you since your abstract is **ACCEPTED** to be presented in the parallel/ round table session of the conference. Authors with accepted abstract(s) are requested to immediatel proceed to payment through BNI bank:

Bank Name: Bank Negara Indonesia, Branch Perguruan Tinggi Bandung, Bandung, INDONESIA
Swift/ BIC: BNINIDJAXXX
Account Number: 0802631015
Account Holder: AASEC

No later than **3 April 2019** and upload the payment proof via website **(NOT EMAIL)** to be able to download LoA and invitation letter.

Thank you and see you in Bali on 24 April 2019.

Don't forget follow our social media accounts and get updates about conference.

Sincerely,
**AASEC Committee**

## AASEC 2019 : Payment Proof for Abstract ABS-902

Eksternal | Kotak Masuk ×

**AASEC 2019** aasec2019.automail@upiconf.org lewat srv148.main-hosting.eu
kepada saya, aasec

Inggris > Indonesia | Terjemahkan pesan

Dear Dr. Mujiarto Mujiarto,

Payment proof for the following abstract has been confirmed.
Payment receipt can be printed directly from your account.

Abstract ID: ABS-902
Title: ANALYSIS OF INJECTOR SPRING DAMAGE TO DETERMINE MAINTENANCE MANAGEMENT DIESEL ENGINE AT PLTD AMPENAN
Author: Wilarso (a*), Tri Wibowo (b), Bambang Teguh (b), Mujiarto (c*)
Presenter: Wilarso
Paid amount: IDR 2300000

Payment receipt can be printed directly from your account as follows:
- Login to your account
- Click "Abstract" on the top menu
- Click "Print Payment Receipt" button below your abstract

Thank you.
Best regards,

AASEC 2019 Organizing Committee
Homepage : http://aasec.conference.upi.edu/2019
Email : aasec@upi.edu

## Reminder (5 days to payment proof submission deadline)

Eksternal | Kotak Masuk ×

**Annual Applied Science Engineering Conference** <aasec@upi.edu>
kepada bcc: saya

Jum, 29 Mar 2019 16.28

Inggris > Indonesia | Terjemahkan pesan | Nonaktifkan untuk: Inggris

Dear participants,

Considering that it is five days to the deadline of payment proof submission deadline, we would like to remind **AUTHORS WITH ACCEPTED ABSTRACTS** to upload their payment proof through their personal accou our submission system **(NOT EMAIL)**.

The payment proof submission system will be closed on **Wednesday, 3 April 2019, 11.59 pm** (there will not be any extended deadline). If you do not upload the payment proof, you are considered **WITHDRAWING** f AASEC 2019.

This is an auto generated email sent to every author; therefore, if you have uploaded the payment proof, please ignore this email.

*Mengingat batas akhir unggah bukti pembayaran tinggal lima hari lagi, kami beritahukan kepada **SELURUH PENULIS YANG TELAH DITERIMA ABSTRAKNYA** untuk segera melakukan pembayaran dan mengung bukti transfer melalui akun masing-masing **(TIDAK MELALUI EMAIL)**.*

*Sistem upload payment akan ditutup pada hari **Rabu 3 April 2019 pukul 23.59 WIB** dan tidak ada perpanjangan waktu. Jika tidak mengunggah bukti transfer pada batas waktu yang telah di tentukan maka saudara/i **dianggap mengundurkan diri**.*

*Email ini dikirim ke seluruh participant, jika saudara/i sudah upload bukti pembayaran mohon untuk mengabaikan email ini.*

Thank you.

## Rundown of AASEC 2019

Eksternal | Kotak Masuk ×

**Annual Applied Science Engineering Conference** <aasec@upi.edu>
kepada bcc: saya

Sab, 13 Apr 2019 16.33

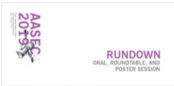
Dear Author(s),

Please find the attached file of AASEC 2019's rundown. We would like you to take a careful look at it since we only put ABS codes on every session. Types of presentations (oral, roundtable, and poster) are the committee's decision. See you in Bali on 24 April 2019.

*Terlampir Susunan Acara pada AASEC 2019. Mohon diperhatikan dan baca dengan seksama. Kami hanya mencantumkan Kode ABS pada setiap sesi. Ketentuan jenis presentasi (oral, roundtable, atau poster) merupakan keputusan panitia yang tidak dapat di ganggu gugat. Sampai jumpa di Bali, 24 April 2019.*

*Terima kasih*
Thank you

Sincerely,
**AASEC Committee**
http://aasec.conference.upi.edu/2019/




RUNDOWN OF AA...

## Full Paper Submission Deadline

Kotak Masuk ×

**Annual Applied Science Engineering Conference** <aasec@upi.edu>
kepada bcc: saya

Sen, 6 Mei 2019 22.1

Inggris > Indonesia Terjemahkan pesan Nonaktifka

Dear author(s),

We would like to remind you that the deadline of full paper submission is on **May 12, 2019, at 11.59 PM (Western Indonesia Standard Time)** and **there will not by any time extension.** Please, submit your full paper(s) through your personal account on submission system (http://aasec.conference.upi.edu/2019/submission/), **DO NOT SEND IT BY EMAIL**.

*If you have already uploaded your full papers via website, just ignore this email and you do not need to re-upload it.*

*Warm regards,*
***Committee***

13 March – Abstract sumbission deadline
20 March – Notification of abstract acceptance
03 April – Payment deadline
24 April – Conference day
12 May – Full paper submission deadline

Noted with thanks. | Thank you for the information. | Thank you for informing me.

## ABS-902; ABS-1066-Revision

**Mujiarto , ST, MT** <mujiarto@umtas.ac.id>
kepada aasec

Rab, 12 Ju

Salam,
Dear Editor AASEC

Berikut revision ABS-902 dan ABS-1066.

Terima kasih.

Wassalam,
Mujiarto

**2 Lampiran** • Dipindai dengan Gmail

ABS-902.docx
2 MB

ABS-1066.doc

Balas | Teruskan

## AASEC Publication Process Update

Kotak Masuk ×

**Annual Applied Science Engineering Conference** <aasec@upi.edu>
kepada bcc: saya

Sen, 12 Agu 2019 16.19

Dear Authors,

We would like to inform you the manuscripts of AASEC 2019 are in publication process by IOP which takes time since they are in a quite long line of other conferences' proceedings publication. We do hope that you are patiently waiting and we will keep you updated. Thank you.

*Diberitahukan kepada para penulis AASEC 2019, bahwa proses penerbitan proceeding masih menunggu antrean dari conference-conference kami yang juga sedang diproses IOP, proses ini akan memakan waktu yang cukup panjang, dimohon untuk bersabar dan menunggu hasil yang akan di informasikan panitia selanjutnya. Terima kasih.*

Warm regards,
**Committee**

Balas    Teruskan

URL JPCS 1402: https://iopscience.iop.org/volume/1742-6596/1402v

URL Pdf: https://iopscience.iop.org/article/10.1088/1742-6596/1402/7/077055/pdf

URL abstract: https://iopscience.iop.org/article/10.1088/1742-6596/1402/7/077055

URL indexing: https://www.scimagojr.com/journalsearch.php?q=130053&tip=sid&clean=0

# Indoor Localization Based WiFi Signal Strength Using Support Vector Machine

**Hani Rubiani**[1*], **Sulidar Fitri**[2], **Muhammad Taufiq**[2], **Mujiarto**[3*]

[1]Department of Electrical Engineering, Universitas Muhammadiyah Tasikmalaya, Indonesia
[2]Department of Information Technology Education, Universitas Muhammadiyah Tasikmalaya, Indonesia
[3]Department of Mechanical Engineering, Universitas Muhammadiyah Tasikmalaya, Indonesia

E-mail: [1]hani.rubiani@umtas.ac.id

**Abstract.**
Estimating the location of Object in an indoor environment poses a fundamental challenge in ubiquitous computing. Indoor localization based on signal strength by utilizing devices in buildings such as WiFi signals is increasingly being done. To determine the user's position using the algorithm of the received signal strength. This paper shows that contrary to popular belief an indoor localization system based on WiFi fingerprints using Support Vector Machine (SVM) method. The performed experiments using 14480 datasets and 302 classes, collected from real world environments in building, and the comparison with Naïve Bayes confirm the effectiveness of SVM-based localization proposal. Experimental results show that the system achieves a correct classification rate of around 88% and minimum average error distance 4.61 meters compared to Naïve Bayes for correct classification rate of around 67% and minimum average error distance 6.21 meters.

## 1. Introduction

The accurate localization of objects and people in indoor environments has long been considered an important building block for ubiquitous computing applications [1,2]. Most research on indoor localization systems has been based on the use of short-range signals, such as WiFi [3,4,5], Bluetooth [6], ultra sound [7], or infrared [8]. GPS (Global Positioning System) very appropriate used in detecting outdoor locations,but less suitable if used in space, because weak or even absence of satellite signals. Because of that, it is necessary to have a stable and accurate system in detecting the location of objects in space, which can be used at home, in the office or in the building others.

This paper shows that contrary to popular belief an indoor localization system based on WiFi fingerprints. With the growth of networks based IEEE 802.11,and increasing variety devices such as laptops, cell phones, and equipment others are WLAN-based, internal location detection space using IEEE 802.11 based technology will growing. Received Signal Strength (RSS) is power radio signal received by the receiver sent by transmitter. In general, RSS will decrease proportional to the distance between the receiver and transmitter [9]. If the relationship between receiver-transmitter distance and signal strength is known, both empirically and analytically, the distance between two devices can be known. There are several advantages to using RSS for indoor localization. First, it can be implemented in a wireless communication system with little even without adding or changing hardware, all that is needed is the ability to obtain and read RSS. Second advantage is no need for synchronization between transmitters and receiver [10]. One important characteristic of RSS is Different orientations provide RSS values different [11]. Different RSS is caused by multipath and also different attenuation.

Support Vector Machine is one method used in classification. SVM is a learning machine method that works on the principle of Structural Risk Minimization (SRM) with the aim of finding the best hyperplane that separates two classes in input space. Support Vector Machines [12, 13] are powerfull techniques used for classification and data regression. They are used for non-parametric supervised classifier for pattern recognition problem.

## 2. Methodology

*A. RSS Based Localization*

Predicting a wireless device's physical location in an indoor environment has been accomplished using techniques based on received signal strength (RSS) [22]-[28], angle of arrival (AoA) [29], [30], time of arrival (ToA) [31], and time difference of arrival (TDoA) [32]. In this paper, we consider only localization techniques that are based on RSS, as these can be constructed with commodity 802.11 hardware and stock drivers.

RSS-based localization refers to the task of estimating an 802.11 device's physical location using only signal strength information. Due to the inherently noise nature of the RSS measurenment, RDD-based localization algorithms typically apply statistical/machine learning techniques, and proceed in two phases :

1. An offline training phase is conducted in which several received signal strength indication (RSSI) readings $\vec{t}_i = (r_{i1}, ........, r_{in})$ are collected over a set of n passive receivers and are labelled with the transmitter's true physical location and orientation $p_i = (x_i, y_i, \theta_i)$ .
2. During the online localization phase, the observed RSSI readings $\vec{o}_j = (r_{j1}, ........, r_{jn})$ are used to produce the device's estimated location $\hat{p}_j = (\hat{x}_j, \hat{y}_j)$ .

*B. Naïve Bayes Classifier*

Localization techniques that use the Naïve Bayes classifier have been proposed in [33], [34], [35], [36] This approach is based on the application of Bayes theorem to obtain a position estimate. Using Bayes theorem, the conditional probability of observing a signal strength vector from the training data at a particular position is computed. During the localization phase, the position estimate is the position that maximizes this probability for the observed signal strength that maximizes this probability for the observed signal strength vector.

Naïve Bayes classifier is a simple probabilistic based on the Bayes theorem in general, Bayes inference specifically with strong (naive) independence assumptions. In the process, Naïve Bayes assumes that the presence or absence of a feature in a class is not related to the presence or absence of other features in the same class.

Generally the attribute $E$ group is represented by a set of attribute values $(AP_1, AP_2, ....AP_n)$ where RSS is the $AP$ attribute value. $L$ is the classification variable in this case, that is the coordinates and l are the values of $L$. From the point of view of opportunities based on the rules of Bayes into class $l$ are:

$$P(l \mid E) = \frac{P(E \mid l)P(l)}{P(E)} \tag{1}$$

To determine the class choice, the maximum opportunities for all $l$ in $L$ are used, with the functions:

$$\underset{l \in L}{\arg\max} \frac{P(E \mid l)P(l)}{P(E)}$$

Because the value is constant for all classes, it can be ignored so as to produce a function:

$$f_l(E) = \underset{l \in L}{\arg\max} \frac{P(E \mid l)P(l)}{P(E)} \tag{2}$$

To overcome various problems, various variants of classifications that use Bayes theorem are proposed, one of which is Naïve Bayes:

$$f_l(E) = \underset{l \in L}{\arg\max}\ P(E \mid l)P(l)\prod_{j=1}^{n} P(E_j \mid l) \tag{3}$$

The approach, called Naïve Bayes, involves modeling signal strength as a Gaussian distribution and using the strength of the signals collected to study Gaussian distribution parameters, which are the mean and standard deviations of training data. As well as calculating the euclidean distance vector $S$ observation signal at position $l$. $S_i$ is the signal strength observed from $AP$ at position $l$, $M_i^l$ is the average signal strength of $AP$ at position $l$ calculated from fingerprint data, $D_i^l$ is the standard deviation of $AP$ in position $l$ calculated from fingerprint data and $|P|$ is the number of $AP$ that reads in position $l$. When the strength of the vector $S$ signal is obtained from the measurement of the current time of the signal strength in the field, the probability $P(S|l)$ is calculated for all positions in the field where the signal strength has been measured during the signal strength database. Position $l$ which has the highest probability $P(S|l)$ for a signal strength vector is classified as the user's position in the field at this time.

$$P(S\,|l)=\prod_{i=1}^{|P|}\left(\frac{1}{\sqrt{2\pi\left(D_i^l\right)^2}}\exp\left(-\frac{\left(S_i-M_i^l\right)^2}{2\left(D_i^l\right)^2}\right)\right) \tag{4}$$

$$M_i^l=\frac{\sum_{i=1}^{n} RSS_i^l}{n} \tag{5}$$

$$D_i^l=\sqrt{\frac{\sum_{i=1}^{n}(RSS_i^l-M_i^l)}{n-1}} \tag{6}$$

*C. Support Vector Machines (SVMs)*

Support Vector Machines [12, 13] are powerfull techniques used for classification and data regression. They are used for non-parametric supervised classifier for pattern recognition problems. SVMs are used in the localization system by training the support vectors on radio map that consist of grid points. SVMs analyze the relationship between the trained fingerprints and their grid points by considering each grid points as a class. The tested RSSI fingerprints are taken as an input to SVM that predict the class to which the tested belongs. This technique can be generalized to classify between more than two classes for $N$ training data $(x_i, y_i)$ .

Before any classification, the RSSI fingerprint vectors are maped into higher dimensional space using kernel function. The SVM kernel function $K(..,..)$ is the dot product of two feature vectors $x_i$ and $x_j$ in some expanded feature space, there are several kernels are proposed by researchers. The four basics kernels as follow : linear, polynomial, sigmoid and radial basis function (RBF). In this research, linear is used in the following form

$$K(x_i, x_j)=x_i^T.x_j \tag{7}$$

where $\sigma^2$ is the variance (i.e width) of the Gaussian kernel.

After representing the training data by mapping the data to the feature space. The SVM algorithms identify hyperplane, which separates the support vector trained with a distance equal $\frac{2}{\|w\|}$. It is constructed in such a way that they can be divided in two data classes with a maximum distance to the closest vector from the same class. The optimization problem is shown in :

$$y_i\,(w^T x_i+b)-1\geq 0 \tag{8}$$

$$\max\frac{2}{\|w\|}\rightarrow\min\frac{\|w\|}{2}\rightarrow\min\frac{1}{2}\|w\|^2. \tag{9}$$

$$\min\left\{L_{pd}\ (w,b,\alpha_i)=\frac{1}{2}\|w\|^2-\sum_{i=1}^{N}\alpha_i\left[y_i\left(w^T x_i+b\right)-1\right]\right\} \tag{10}$$

Where $b$ and $\alpha_i$ are solution of the contrains and $y_i$ is the output of each class $y_i \in \{1,-1\}$, which achieve the minimize (9) based on lagrangian function, where $\alpha_i$ is the lagrangian multipliers. The constrained optimization problem can be expressed in a dual form by searching a solution under the form [38].

$$w = \sum_{i=1}^{N} \alpha_i y_i x_i \qquad \sum_{i=1}^{N} \alpha_i y_i = 0 \tag{11}$$

Maximizing with respect to $\alpha$ :

$$\max \left\{ \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i . x_j^T) \right\} \tag{12}$$

Under constrains :

$$\sum_{i=1}^{N} \alpha_i y_i = 0 \text{ where } \alpha_i \geq 0 \quad \forall i \tag{13}$$

There is a pure decision and uncertainty is inefficient. A pure decision means each node's data subset contains one and only one target location.

## 3. Experimental Result

The research material used indoor localization based WiFi using Support Vector Machine method is the result of the measurenment of signal strength received by laptop for IEEE 802.11g. The experiment was carried out in the corridor of the 3rd floor of the building with an area of ± 302 m2. Determination of reference points is the next stage in space planning which is the scope of research. At this stage the corridor is measured and then divided into areas with an area of 1 $m^2$. Before taking data training, the reference points that are right in the middle of each area are marked first and it is certain that the mark is in accordance with the coordinates that will be used as classifications in data fingerprint, this is done to facilitate the process of measuring data training.

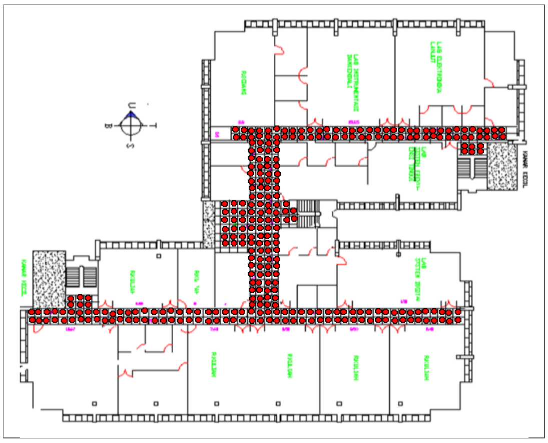


**Figure 1. Reference Points of 1 $m^2$**

This research is divided into several stages. These stages are :

1. Space Planning is the first step in making a signal strength map in that space become the scope of research. The research room is floor hall 1,2, and 3 a building. At this stage the aisle is measured then divided into cells with a width of 1 meter each cell.
2. Measurement of RSS. In this process RSS measurements are received by the Laptop, at each cell that has been measured. Measurement done in the same direction ie West. Measurement RSS is done when the building conditions are quiet, that is on Saturdays and Sundays or holidays. Software the one used is NetSurveyor.

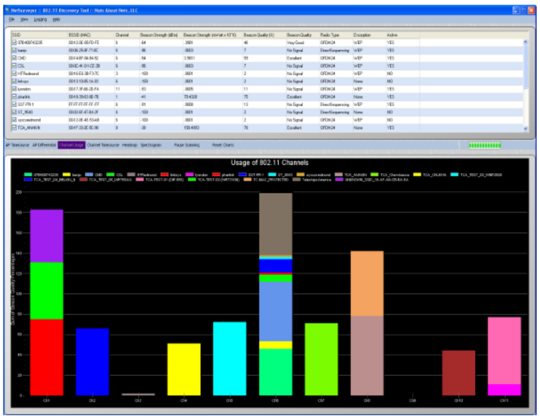



**Figure 2. NetSurveyor View**

3. RSS visualization. RSS visualization is used for provide a map of the signal strength received (RSS). RSS value on each grid point is obtained by calculating the average the signal received on the grid point. Visualization is done using software RapidMiner. Visualization using the AP installed on the 3rd floor of the building.

**Figure 3. RSS Visualization for AP1**

4. Data Validation
   Dividing the dataset for validation and for training, need to know whether the model to be made is the best model. Then finally use the statistical method to estimate the accuracy of the model against previously unknown data.

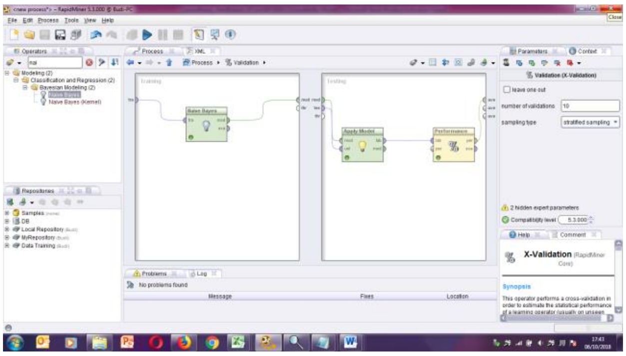



**Figure 4. Data Validation using RapidMiner**

5. Location Fingerprinting. Fingerprinting is method for measurable data mapping, namely RSS to in a grid-point that covers the entire localization area which will be used for location estimation. RSS not all of the

AP in the building environment used, but RSS is chosen from several APs has a significant influence on location estimation. RSS data processing is done using Excel software.

6. Algorithm modeling. This process is carried out algorithm modeling used is Naïve Bayes and Support Vector Machine. In the process this modeling can be known the percentage accuracy.

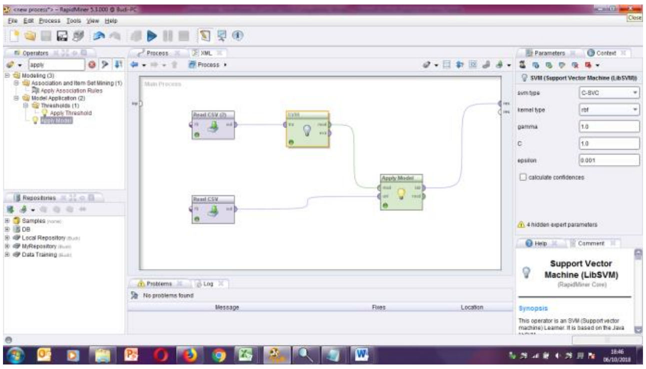


**Figure 5. Algorithm Modeling**

7. Location Tests and Estimates. Testing is an important process for knowing results from a system. Test data obtained by how to do RSS measurements received laptop by walking along the 1st floor hallway, 2 and buildings. Estimated location of objects in this case is an IEEE 802.11 based laptop, obtained from a comparison between RSS measurements in fact, the test data with previous measurements that have been stored in fingerprint. Location estimation using algorithm which has been modeled before. Estimated error location is obtained by calculating distance between the actual location and the estimated location.

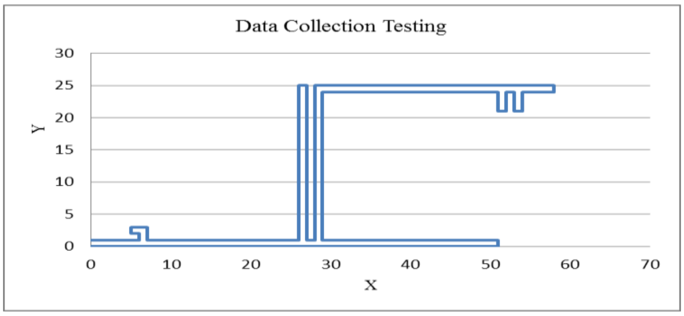
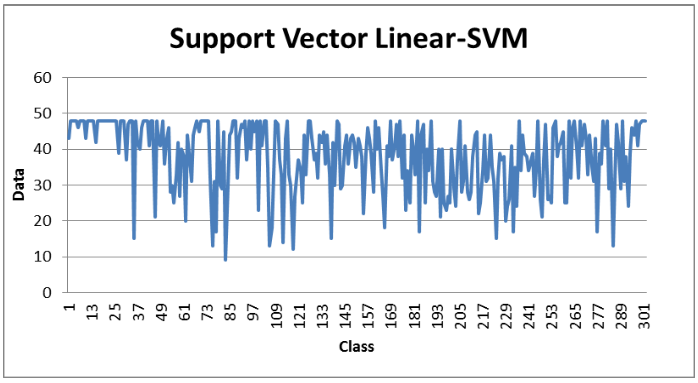



**Figure 6. Data Collection Testing**

8. Analysis of results. At this stage an analysis is carried out for knowing the magnitude of the estimated error on algorithm used.

## 4. Evaluation

Hasil penelitian




| Algorithm | Error (m) |
|---|---|
| Naïve Bayes | 6,21 |
| Linear SVM (C=5) | 5,884873 |
| Linear SVM (C=1) | 4,612462 |
| Linear SVM (C=2) | 4,612462 |

## 6. Conclusion

In this study,

## References


[1] A. Hopper, A. Harter, and T. Blackie. The active badge system. InProc. Of INTERCHI-93, pages 533–534, Amsterdam, The Netherlands, 1993.

[2] T. Kindberg and A. Fox. System software for ubiquitous computing.IEEE Pervasive Computing, 1(1):26–35, Jan. 2002.

[3] P. Bahl and V. N. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. In INFOCOM, pages 775–784, 2000.

[4] E. Elnahrawy, X. Li, and R. Martin. The limits of localization using signal strength: A comparative study. In Proceedings of the 1st IEEE International Conference on Sensor and Ad Hoc Communications and Networks, Santa Clara, CA, Oct. 2004.

[5] A. Ladd, K. Bekris, G. Marceau, A. Rudys, L. Kavraki, and D. Wallach. Robotics-based location sensing using wireless ethernet. In Proceedings of the Tenth ACM International Conference on Mobile Computing and Networking (MOBICOM), 2002.

[6] L. Aalto, N. Gothlin, J. Korhonen, and T. Ojala. Bluetooth and WAP push based location-aware mobile advertising system. In MobiSYS ’04: Proceedings of the 2nd international conference on Mobile systems, applications, and services, pages 49–58. ACM Press, 2004.

[7] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan. The cricket location-support system. In Mobile Computing and Networking, pages 32–43, 2000.

[8] 16. A. Ward, A. Jones, and A. Hopper. A new location technique for the active office. In IEEE Personnel Communications, 4(5), pages 42–47, Oct. 1997.

[9] A. Kupper, 2005. Location Based Services: Fundamentals and Operations, John Wiley & Sons.

[10] J. Caffery, 2000. Wireless Location in CDMA Cellular Radio Systems, Kluwer Academic Publishers.

[11] K. Kamol and K. Prashant, 2004. Properties of Indoor Received Signal Strength for WLAN Location Fingerprinting, in: *Proceedings of the First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Service.*

[12] Schölkopf B. and Smola A., "A. Learning with Kernels, " MIT Press, Cambridge, MA, 2002.

[13] T Joachims, "Transductive inference for text classification using support vector machines, " in Proceeding of International Conference on Machine Learning (ICML), pp. 200–209, 1999.
[14] N. Cristianini and J. Shawe-Taylor, “An introduction to support vector machines and other kernel-based learning methods”, Cambridge University press, (2000).
[15] V. Kecman, “Learning and soft computing: support vector machines, neural networks, and fuzzy logic models”, MIT press, (2001).
[16] V. Vapnik, “The nature of statistical learning Theroy’, Springer (2000).
[17] F. Yu, M. Jiang, J. Liang , X. Qiao, T. Peng and X. Hu. “An Improved Indoor Localization of WiFi Based on Support Vector Machines’, International Journal of Future Generation Communication and Networking Vol.7, No. 5 pp. 191-206 (2014).
[18] S. Ke, M. Zheniie, Z. Rentong, H. Wenbiao and C. Hongsheng, “Support Vector Regression Based Indoor location in IEEE 802.11 Environments”, Mobile Information Systems Vol. 2015, Article ID 295652, 14 pages (2015).
[19] Z. A. Deng, Y. B. Xu, and D. Wu, “Indoor positioning in WLAN environment based on support vector regression and space partitioning,” Advanced Materials Research, vol. 204–210, pp. 1599–1602, 2011.
[20] Z. Yu, D. Lian, L. Lei and H. Lizhi,” Study of Indoor Positioning Method Based on Combination of Support Vector Regression and Kalman Filtering “, International Journal of Future Generation Communication and Networking Vol. 9, No. 3, pp. 201 – 214 (2016)
[21] Z.L. Wu, C.H. Li, J.Y. Ngand K. Leung, “Location Estimation via Support Vector Regression”, IEEE Transactions on Mobile Computer, vol. 6, no. 3, (2007), pp.311–321.
[22] P. Bahl and V. N. Padmanabhan, “RADAR: An in-building RF-based user location and tracking system,” in *INFOCOM*,2000,pp.775–784.
[23] D. Madigan, E. Elnahrawy, R. P. Martin, W.-H. Ju, P. Krishnan, and A. Krishnakumar, “Bayesian indoor positioning systems.” in *INFOCOM*. IEEE, 2005, pp. 1217–1227.
[24] J. Ash and L. Potter, “Sensor network localization via received signal strength measurements with directional antennas,” *Conference on Communication, Control, and Computing*,pp.1861–1870,Sept2004.
[25] A. M. Ladd, K. E. Bekris, A. Rudys, L. E. Kavraki, and D. S. Wallach,“Robotics-based location sensing using wireless ethernet,” in *MobiCom*, 2002, pp. 227–238.
[26] T. King, S. Kopf, T. Haenselmann, C. Lubberger, and W. Effelsberg,“Compass: A probabilistic indoor positioning system based on 802.11and digital compasses,” in *WiNTECH*,2006,pp.34–40.
[27] M. A. Youssef, A. Agrawala, and A. U. Shankar, “Wlan location determination via clustering and probability distributions,” in *PERCOM*, Washington, DC, USA, 2003.
[28] E. Elnahrawy, X. Li, and R. P. Martin, “The limits of localization usin signal strength: A comparative study,” in *SECON*.IEEE,2004.
[29] E. Elnahrawy, J. Austen-Francisco, and R. P. Martin, “Adding angle of arrival modality to basic RSS location management techniques,” in *ISWPC*,2007,pp.464–469.
[30] D. Niculescu and B. Nath, “VOR base stations for indoor 802.11 positioning,” in *MobiCom*,2004,pp.58–69.
[31] Y.-T. Chan, W.-Y. Tsui, H.-C. So, and P. chung Ching, “Time-of-arrival based localization under NLOS conditions,” *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY*,vol.55,no.1,January2006.
[32] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan, “The cricket location-support system,” in *MOBICOM*,2000.
[33] P. Tao, A. Rudys, A. M. Ladd, and D. S. Wallach, “Wireless LAN location-sensing for security applications,” in *WiSe*,2003.
[34] A. M. Ladd, K. E. Bekris, A. Rudys, L. E. Kavraki, and D. S. Wallach, “Robotics-based location sensing using wireless ethernet,” in *MobiCom*, 2002, pp. 227–238.
[35] E. Elnahrawy, X. Li, and R. P. Martin, “The limits of localization using signal strength: A comparative study,” in *SECON*.IEEE,2004.
[36] A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach, and L. E. Kavraki, “Practical robust localization over large-scale 802.11 wireless networks,” in *MobiCom*,2004,pp.70–84.

[38] M. Brunato and R. Battiti, "Statistical learning theory for location fingerprinting in wireless LANs, " Comput. Netw. 47, 825–845, 2005.

# Indoor Localization Based WiFi Signal Strength Using Support Vector Machine

**Hani Rubiani**[1*], **Sulidar Fitri**[2], **Muhammad Taufiq**[2], **Mujiarto**[3*]

[1]Department of Electrical Engineering, Universitas Muhammadiyah Tasikmalaya, Indonesia
[2]Department of Information Technology Education, Universitas Muhammadiyah Tasikmalaya, Indonesia
[3]Department of Mechanical Engineering, Universitas Muhammadiyah Tasikmalaya, Indonesia

*hani.rubiani@umtas.ac.id; mujiarto@umtas.ac.id

**Abstract.** Ubiquitous computing to estimate the location of objects in a building raises a fundamental challenge and there has been a lot of research on localization in buildings based on signal strength by utilizing devices inside such as Wifi signals. Positioning objects using algorithms of received signal strength in this paper using Linear Support Vector Machine which will be compared with Naïve Bayes. Experiments carried out using 14480 datasets and 302 classes were collected from the real world environment and the results showed that the system reached the correct classification level of around 88% and a minimum distance of error of 4.61 meters compared to Naïve Bayes for the correct classification level of around 67 % and average error distance of 6.21 meters.

## 1. Introduction

The accurate localization of objects and people in indoor environments has long been considered an important building block for ubiquitous computing applications[1][2]. Most research on indoor localization systems has been based on the use of short-range signals, such as WiFi [3]–[5]Bluetooth [6], ultra sound[7], or infrared [8]. GPS (Global Positioning System) very appropriate used in detecting outdoor locations, but less suitable if used in space, because weak or even absence of satellite signals. Because of that, it is necessary to have a stable and accurate system in detecting the location of objects in space, which can be used at home, in the office or in the building others.

This paper shows that contrary to popular belief an indoor localization system based on WiFi fingerprints. With the growth of networks based IEEE 802.11,and increasing variety devices such as laptops, cell phones, and equipment others are WLAN-based, internal location detection space using IEEE 802.11 based technology will growing. Received Signal Strength (RSS) is power radio signal received by the receiver sent by transmitter. In general, RSS will decrease proportional to the distance between the receiver and transmitter [9]. If the relationship between receiver-transmitter distance and signal strength is known, both empirically and analytically, the distance between two devices can be known. There are several advantages to using RSS for indoor localization. First, it can be implemented in a wireless communication system with little even without adding or changing hardware, all that is needed is the ability to obtain and read RSS. Second advantage is no need for synchronization between transmitters and receiver[10]. One important characteristic of RSS is Different orientations provide RSS values different [11]. Different RSS is caused by multipath and also different attenuation.

Support Vector Machine is one method used in classification. SVM is a learning machine method that works on the principle of Structural Risk Minimization (SRM) with the aim of finding the best hyperplane that separates two classes in input space. Support Vector Machines [12], [13] are powerful techniques used for classification and data regression. They are used for non-parametric supervised classifier for pattern recognition problem.

## 2. Methods

### *2.1. RSS Based Localization*

Predicting a wireless device's physical location in an indoor environment has been accomplished using techniques based on received signal strength (RSS) [3]–[5], [14]–[17],angle of arrival (AoA) [18],[19], time of arrival (ToA) [20],and time difference of arrival (TDoA) [7]. In this paper, we consider only localization techniques that are based on RSS, as these can be constructed with commodity 802.11 hardware and stock drivers.

RSS-based localization refers to the task of estimating an 802.11 device's physical location using only signal strength information. Due to the inherently noise nature of the RSS measurement, RDD-based localization algorithms typically apply statistical/machine learning techniques, and proceed in two phases :

1. An offline training phase is conducted in which several received signal strength indication (RSSI) readings $\vec{t}_i = (r_{i1},......., r_{in})$ are collected over a set of n passive receivers and are labelled with the transmitter's true physical location and orientation $p_i = (x_i, y_i, \theta_i)$ .
2. During the online localization phase, the observed RSSI readings $\vec{o}_j = (r_{j1},......., r_{jn})$ are used to produce the device's estimated location $\hat{p}_j = (\hat{x}_j, \hat{y}_j)$ .

### *2.2. Naïve Bayes Classifier*

Localization techniques that use the Naïve Bayes classifier have been proposed in [33], [34], [35], [36] This approach is based on the application of Bayes theorem to obtain a position estimate. Using Bayes theorem, the conditional probability of observing a signal strength vector from the training data at a particular position is computed. During the localization phase, the position estimate is the position that maximizes this probability for the observed signal strength that maximizes this probability for the observed signal strength vector.

Naïve Bayes classifier is a simple probabilistic based on the Bayes theorem in general, Bayes inference specifically with strong (naive) independence assumptions. In the process, Naïve Bayes assumes that the presence or absence of a feature in a class is not related to the presence or absence of other features in the same class.

Generally the attribute $E$ group is represented by a set of attribute values $(AP_1, AP_2, ....AP_n)$ where RSS is the $AP$ attribute value. $L$ is the classification variable in this case, that is the coordinates and l are the values of $L$. From the point of view of opportunities based on the rules of Bayes into class $l$ are:

$$P(l \mid E) = \frac{P(E \mid l)P(l)}{P(E)} \tag{1}$$

To determine the class choice, the maximum opportunities for all $l$ in $L$ are used, with the functions:

$$\underset{l \in L}{\arg\max} \frac{P(E \mid l)P(l)}{P(E)}$$

Because the value is constant for all classes, it can be ignored so as to produce a function:

$$f_l(E) = \underset{l \in L}{\arg\max} \frac{P(E \mid l)P(l)}{P(E)} \tag{2}$$

To overcome various problems, various variants of classifications that use Bayes theorem are proposed, one of which is Naïve Bayes:

$$f_l(E) = \underset{l \in L}{\arg\max} \; P(E \mid l)P(l)\prod_{j=1}^{n} P(E_j \mid l) \tag{3}$$

The approach, called Naïve Bayes, involves modeling signal strength as a Gaussian distribution and using the strength of the signals collected to study Gaussian distribution parameters, which are the mean and standard deviations of training data. As well as calculating the Euclidean distance vector $S$ observation signal at position $l$ . $S_i$ is the signal strength observed from $AP$ at position $l$ , $M_i^l$ is the average signal strength of $AP$ at position $l$ calculated from fingerprint data, $D_i^l$ is the standard deviation of $AP$ in position $l$ calculated from fingerprint data and $|P|$ is the number of $AP$ that reads in position $l$ . When the strength of the vector $S$ signal is obtained from the measurement of the current time of the signal strength in the field, the probability $P(S|l)$ is calculated for all positions in the field where the signal strength has been measured during the signal strength database. Position $l$ which has the highest probability $P(S|l)$ for a signal strength vector is classified as the user's position in the field at this time.

$$P(S|l) = \prod_{i=1}^{|P|} \left( \frac{1}{\sqrt{2\pi \left( D_i^l \right)^2}} \exp \left( - \frac{\left( S_i - M_i^l \right)^2}{2\left( D_i^l \right)^2} \right) \right) \tag{4}$$

$$M_i^l = \frac{\sum_{i=1}^{n} RSS_i^l}{n} \tag{5}$$

$$D_i^l = \sqrt{\frac{\sum_{i=1}^{n} (RSS_i^l - M_i^l)}{n-1}} \tag{6}$$

*2.3. Support Vector Machines (SVMs)*

Support Vector Machines [12], [13] are powerfull techniques used for classification and data regression. They are used for non-parametric supervised classifier for pattern recognition problems. SVMs are used in the localization system by training the support vectors on radio map that consist of grid points. SVMs analyze the relationship between the trained fingerprints and their grid points by considering each grid points as a class. The tested RSSI fingerprints are taken as an input to SVM that predict the class to which the tested belongs. This technique can be generalized to classify between more than two classes for $N$ training data $(x_i, y_i)$ .

Before any classification, the RSSI fingerprint vectors are mapped into higher dimensional space using kernel function. The SVM kernel function $K(..,..)$ is the dot product of two feature vectors $x_i$ and $x_j$ in some expanded feature space, there are several kernels are proposed by researchers. The four basics kernels as follow: linear, polynomial, sigmoid and radial basis function (RBF). In this research, linear is used in the following form

$$K(x_i, x_j) = x_i^T . x_j \tag{7}$$

where $\sigma^2$ is the variance (i.e width) of the Gaussian kernel.

After representing the training data by mapping the data to the feature space. The SVM algorithms identify hyperplane, which separates the support vector trained with a distance equal $\frac{2}{\|w\|}$. It is constructed in such a way that they can be divided in two data classes with a maximum distance to the closest vector from the same class. The optimization problem is shown in :

$$y_i (w^T x_i + b) - 1 \geq 0 \tag{8}$$

$$\max \frac{2}{\|w\|} \to \min \frac{\|w\|}{2} \to \min \frac{1}{2}\|w\|^2 . \tag{9}$$

$$\min \left\{ L_{pd} \ (w,b,\alpha_i) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \alpha_i \left[ y_i \left( w^T \ x_i + b \right) - 1 \right] \right\} \tag{10}$$

Where $b$ and $\alpha_i$ are solution of the constrains and $y_i$ is the output of each class $y_i \in \{1,-1\}$, which achieve the minimize (9) based on lagrangian function, where $\alpha_i$ is the lagrangian multipliers. The constrained optimization problem can be expressed in a dual form by searching a solution under the form [38].

$$w = \sum_{i=1}^{N} \alpha_i y_i x_i \qquad \sum_{i=1}^{N} \alpha_i y_i = 0 \tag{11}$$

Maximizing with respect to $\alpha$ :

$$\max \left\{ \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \, \alpha_j \; y_i \; y_j \; (x_i \, . x_j^T \,) \right\} \tag{12}$$

Under constrains :

$$\sum_{i=1}^{N} \alpha_i y_i = 0 \text{ where } \alpha_{i \geq 0} \quad \forall i \tag{13}$$

There is a pure decision and uncertainty is inefficient. A pure decision means each node's data subset contains one and only one target location.

### 3. Results and Discussion

The research material used indoor localization based WiFi using Support Vector Machine method is the result of the measurement of signal strength received by laptop for IEEE 802.11g. The experiment was carried out in the corridor of the 3rd floor of the building with an area of ± 302 m2. Determination of reference points is the next stage in space planning which is the scope of research. At this stage the corridor is measured and then divided into areas with an area of 1 m$^2$. Before taking data training, the reference points that are right in the middle of each area are marked first and it is certain that the mark is in accordance with the coordinates that will be used as classifications in data fingerprint, this is done to facilitate the process of measuring data training.

**Figure 1.** Reference Points of 1 m$^2$

This research is divided into several stages. These stages are :

1. Space Planning is the first step in making a signal strength map in that space become the scope of research. The research room is floor hall 1,2, and 3 a building. At this stage the aisle is measured then divided into cells with a width of 1 meter each cell.
2. Measurement of RSS. In this process RSS measurements are received by the Laptop, at each cell that has been measured. Measurement done in the same direction ie West. Measurement RSS is done when the building conditions are quiet, that is on Saturdays and Sundays or holidays. Software the one used is NetSurveyor.

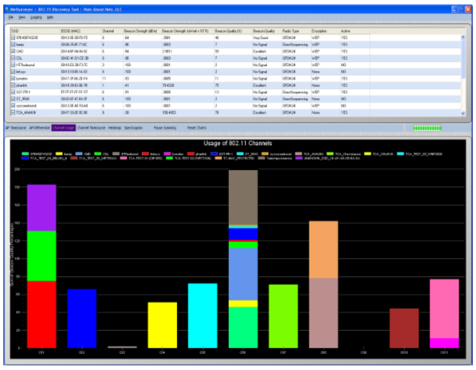


**Figure 2.** NetSurveyor View

3. RSS visualization. RSS visualization is used for provide a map of the signal strength received (RSS). RSS value on each grid point is obtained by calculating the average the signal received on the grid point. Visualization is done using software RapidMiner. Visualization using the AP installed on the 3rd floor of the building.

**Figure 3.** RSS Visualization for AP1

4. Data Validation
   Dividing the dataset for validation and for training, need to know whether the model to be made is the best model. Then finally use the statistical method to estimate the accuracy of the model against previously unknown data.

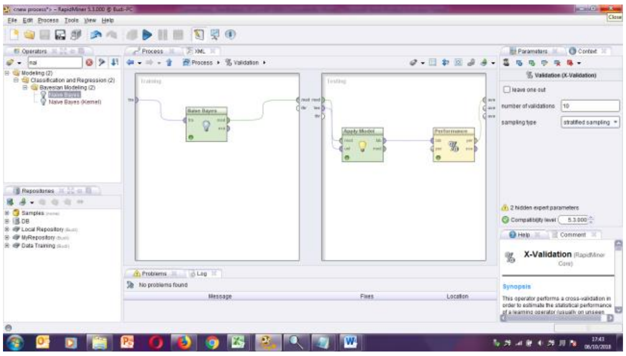


**Figure 4.** Data Validation using RapidMiner

5. Location Fingerprinting. Fingerprinting is method for measurable data mapping, namely RSS to in a grid-point that covers the entire localization area which will be used for location estimation. RSS not all of the AP in the building environment used, but RSS is chosen from several APs has a significant influence on location estimation. RSS data processing is done using Excel software.
6. Algorithm modeling. This process is carried out algorithm modeling used is Naïve Bayes and Support Vector Machine. In the process this modeling can be known the percentage accuracy.

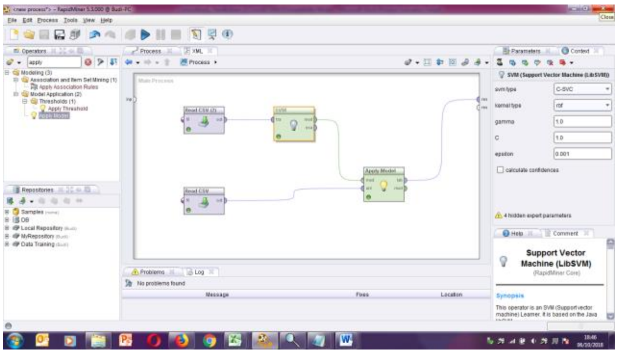


**Figure 5. Algorithm Modeling**

7. Location Tests and Estimates. Testing is an important process for knowing results from a system. Test data obtained by how to do RSS measurements received laptop by walking along the 1st floor hallway, 2 and buildings. Estimated location of objects in this case is an IEEE 802.11 based laptop, obtained from a comparison between RSS measurements in fact, the test data with previous measurements that have been stored in fingerprint. Location estimation using algorithm which has been modeled before. Estimated error location is obtained by calculating distance between the actual location and the estimated location.

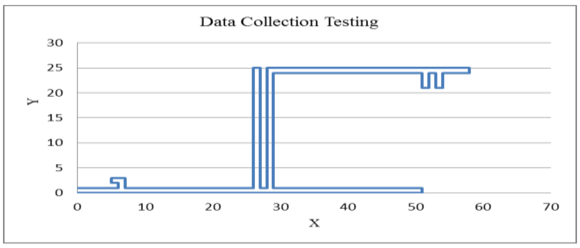



**Figure 6. Data Collection Testing**

8. Analysis of results. At this stage an analysis is carried out for knowing the magnitude of the estimated error on algorithm used.

Results of research in figure 7 shows the Class for indoor collection testing data for localization based wifi signal using a support vector linear machine. Data testing was collected in the number of 10 to 50 data taken by walking along the hallway area in the building.

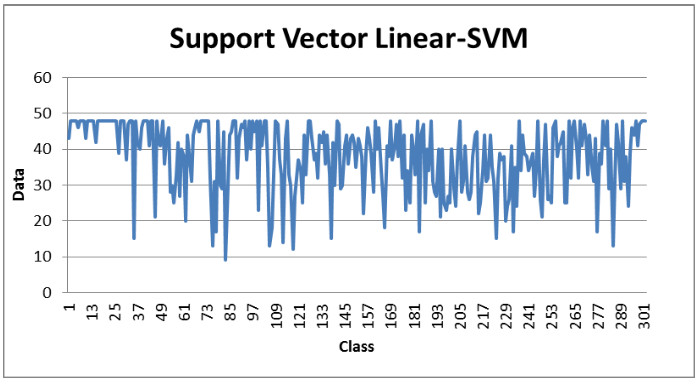



**Figure 7.** Class of Testing Data

Table 1 shows the results of the minimum distance error by comparing the differences in the methods used namely Naïve Bayes and Linear SVM with C = 1, 2 and 5.

**Table 1.** Comparison of the minimum average distance with Niave Bayes and Linear SVM methods

| Algorithm | Error (m) |
|---|---|
| Naïve Bayes | 6,21 |
| Linear SVM (C=5) | 5,884873 |
| Linear SVM (C=1) | 4,612462 |
| Linear SVM (C=2) | 4,612462 |

## 4. Conclusions

The SVM Linear method is better in the results of the minimum average error distance compared to Naïve Bayes. Linear SVM with C = 1, 2 and 5 shows that C = 1 and 2 get a fixed result but when C = 5 shows the change in the results of the minimum average error distance is not good compared to C = 1 and 2.

## References

[1] B. A, Hopper. A, Harter. T, "The active badge system," *InProc. Of INTERCHI-93*, pp. 533–534, 1993.
[2] F. T, Kindberg . A, "System software for ubiquitous computing," *IEEE Pervasive Comput.*, vol. 1, no. 1, pp. 26–35, 2002.
[3] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," *INFOCOM*, pp. 775–784, 2000.
[4] and R. M. E. Elnahrawy, X. Li, "The limits of localization using signal strength: A comparative study.," *Proc. 1st IEEE Int. Conf. Sens. Ad Hoc Commun. Networks, St. Clara, CA*, 2004.
[5] and D. W. A. Ladd, K. Bekris, G. Marceau, A. Rudys, L. Kavraki, "Robotics-based location sensing using wireless ethernet," *Proc. Tenth ACM Int. Conf. Mob. Comput. Netw.*, 2002.
[6] and T. O. L. Aalto, N. Gothlin, J. Korhonen, "Bluetooth and WAP push based location-aware mobile advertising system. In MobiSYS '04," *Proc. 2nd Int. Conf. Mob. Syst. Appl. Serv.*, pp. 49–58, 2004.
[7] and H. B. N. B. Priyantha, A. Chakraborty, "The cricket location-support system," *Mob. Comput. Netw.*, pp. 32–43, 2000.
[8] and A. H. A. Ward, A. Jones, "A new location technique for the active office," *IEEE Pers. Commun.*, vol. 4, no. 5, pp. 42–47, 1997.

[9] A. Kupper, *Location Based Services: Fundamentals and Operations*. John Wiley & Sons, 2005.
[10] J. Caffery, *Wireless Location in CDMA Cellular Radio Systems*. Kluwer Academic Publishers, 2000.
[11] K. Kamol and K. Prashant, "Properties of Indoor Received Signal Strength for WLAN Location Fingerprinting," *Proc. First Annu. Int. Conf. Mob. Ubiquitous Syst. Netw. Serv.*, 2004.
[12] Schölkopf B. and Smola A., *A. Learning with Kernels*. MIT Press, Cambridge, MA, 2002.
[13] T Joachims, "Transductive inference for text classification using support vector machines," *Proceeding Int. Conf. Mach. Learn.*, pp. 200–209, 1999.
[14] and A. K. D. Madigan, E. Elnahrawy, R. P. Martin, W.-H. Ju, P. Krishnan, "Bayesian indoor positioning systems," *INFOCOM. IEEE*, pp. 1217–1227, 2005.
[15] J. Ash and L. Potter, "Sensor network localization via received signal strength measurements with directional antennas," *Conf. Commun. Control. Comput.*, pp. 1861–1870, 2004.
[16] and W. E. T. King, S. Kopf, T. Haenselmann, C. Lubberger, "Compass: A probabilistic indoor positioning system based on 802.11and digital compasses," *WiNTECH*, pp. 34–40, 2006.
[17] and A. U. S. M. A. Youssef, A. Agrawala, *Wlan location determination via clustering and probability distributions*. Washington, DC, USA, 2003.
[18] and R. P. M. E. Elnahrawy, J. Austen-Francisco, "Adding angle of arrival modality to basic RSS location management techniques," *ISWPC*, pp. 464–469, 2007.
[19] D. Niculescu and B. Nath, "VOR base stations for indoor 802.11 positioning," *MobiCom*, pp. 58–69, 2004.
[20] and P. chung C. Y.-T. Chan, W.-Y. Tsui, H.-C. So, "Time-of-arrival based localization under NLOS conditions," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, 2006.

PAPER • OPEN ACCESS

# Indoor localization based Wi-Fi signal strength using support vector machine

To cite this article: H Rubiani *et al* 2019 *J. Phys.: Conf. Ser.* **1402** 077055

View the article online for updates and enhancements.

# Indoor localization based Wi-Fi signal strength using support vector machine

**H Rubiani[1,*], S Fitri[2], M Taufiq[2] and M Mujiarto[3]**

[1] Department of Electrical Engineering, Universitas Muhammadiyah Tasikmalaya, Indonesia
[2] Department of Information Technology Education, Universitas Muhammadiyah Tasikmalaya, Indonesia
[3] Department of Mechanical Engineering, Universitas Muhammadiyah Tasikmalaya, Indonesia

*hani.rubiani@umtas.ac.id

**Abstract**. Ubiquitous computing to estimate the location of objects in a building raises a fundamental challenge and there has been a lot of research on localization in buildings based on signal strength by utilizing devices inside such as Wi-Fi signals. Positioning objects using algorithms of received signal strength in this paper using Linear Support Vector Machine which will be compared with Naïve Bayes. Experiments carried out using 14480 datasets and 302 classes were collected from the real world environment and the results showed that the system reached the correct classification level of around 88% and a minimum distance of error of 4.61 meters compared to Naïve Bayes for the correct classification level of around 67 % and average error distance of 6.21 meters.

## 1. Introduction

The accurate localization of objects and people in indoor environments has long been considered an important building block for ubiquitous computing applications[1,2]. Most research on indoor localization systems has been based on the use of short-range signals, such as Wi-Fi [3–5], Bluetooth [6], ultra sound [7], or infrared [8]. GPS (Global Positioning System) very appropriate used in detecting outdoor locations, but less suitable if used in space, because weak or even absence of satellite signals. Because of that, it is necessary to have a stable and accurate system in detecting the location of objects in space, which can be used at home, in the office or in the building others.

This paper shows that contrary to popular belief an indoor localization system based on Wi-Fi fingerprints. With the growth of networks based IEEE 802.11,and increasing variety devices such as laptops, cell phones, and equipment others are WLAN-based, internal location detection space using IEEE 802.11 based technology will growing. Received Signal Strength (RSS) is power radio signal received by the receiver sent by transmitter. In general, RSS will decrease proportional to the distance between the receiver and transmitter [9]. If the relationship between receiver-transmitter distance and signal strength is known, both empirically and analytically, the distance between two devices can be known. There are several advantages to using RSS for indoor localization. First, it can be implemented in a wireless communication system with little even without adding or changing hardware, all that is needed is the ability to obtain and read RSS. Second advantage is no need for synchronization between transmitters and receiver [10]. One important characteristic of RSS is Different orientations provide RSS values different [11]. Different RSS is caused by multipath and also different attenuation.

Support Vector Machine is one method used in classification. SVM is a learning machine method that works on the principle of Structural Risk Minimization (SRM) with the aim of finding the best hyperplane that separates two classes in input space. Support Vector Machines are powerful techniques used for classification and data regression [12,13]. They are used for non-parametric supervised classifier for pattern recognition problem.

## 2. Methods

### *2.1. RSS based localization*

Predicting a wireless device's physical location in an indoor environment has been accomplished using techniques based on received signal strength (RSS) [3–5,14-17] angle of arrival (AoA) [18,19] time of arrival (ToA) [20] and time difference of arrival (TDoA) [7]. In this paper, we consider only localization techniques that are based on RSS, as these can be constructed with commodity 802.11 hardware and stock drivers.

RSS-based localization refers to the task of estimating an 802.11 device's physical location using only signal strength information. Due to the inherently noise nature of the RSS measurement, RDD-based localization algorithms typically apply statistical/machine learning techniques, and proceed in two phases:

- An offline training phase is conducted in which several received signal strength indication (RSSI) readings $\vec{t}_i = (r_{i1},........,r_{in})$ are collected over a set of n passive receivers and are labelled with the transmitter's true physical location and orientation $p_i = (x_i, y_i, \theta_i)$ .
- During the online localization phase, the observed RSSI readings $\vec{o}_j = (r_{j1},........,r_{jn})$ are used to produce the device's estimated location $\hat{p}_j = (\hat{x}_j, \hat{y}_j)$ .

### *2.2. Naïve Bayes classifier*

Localization techniques that use the Naïve Bayes classifier have been proposed in [33-36]. This approach is based on the application of Bayes theorem to obtain a position estimate. Using Bayes theorem, the conditional probability of observing a signal strength vector from the training data at a particular position is computed. During the localization phase, the position estimate is the position that maximizes this probability for the observed signal strength that maximizes this probability for the observed signal strength vector.

Naïve Bayes classifier is a simple probabilistic based on the Bayes theorem in general, Bayes inference specifically with strong (naive) independence assumptions. In the process, Naïve Bayes assumes that the presence or absence of a feature in a class is not related to the presence or absence of other features in the same class.

Generally the attribute $E$ group is represented by a set of attribute values $(AP_1, AP_2, ....AP_n)$ where RSS is the $AP$ attribute value. $L$ is the classification variable in this case, that is the coordinates and l are the values of $L$ . From the point of view of opportunities based on the rules of Bayes into class $l$ are:

$$P(l \mid E) = \frac{P(E \mid l)P(l)}{P(E)} \tag{1}$$

To determine the class choice, the maximum opportunities for all $l$ in $L$ are used, with the functions:

$$\arg\max_{l \in L} \frac{P(E \mid l)P(l)}{P(E)}$$

Because the value is constant for all classes, it can be ignored so as to produce a function:

$$f_l(E) = \underset{l \in L}{\arg\max} \frac{P(E \mid l)P(l)}{P(E)} \tag{2}$$

To overcome various problems, various variants of classifications that use Bayes theorem are proposed, one of which is Naïve Bayes:

$$f_l(E) = \underset{l \in L}{\arg\max} \; P(E \mid l)P(l)\prod_{j=1}^{n} P(E_j \mid l) \tag{3}$$

The approach, called Naïve Bayes, involves modelling signal strength as a Gaussian distribution and using the strength of the signals collected to study Gaussian distribution parameters, which are the mean and standard deviations of training data. As well as calculating the Euclidean distance vector $S$ observation signal at position $l$. $S_i$ is the signal strength observed from $AP$ at position $l$, $M_i^l$ is the average signal strength of $AP$ at position $l$ calculated from fingerprint data, $D_i^l$ is the standard deviation of $AP$ in position $l$ calculated from fingerprint data and $|P|$ is the number of $AP$ that reads in position $l$. When the strength of the vector $S$ signal is obtained from the measurement of the current time of the signal strength in the field, the probability $P(S|l)$ is calculated for all positions in the field where the signal strength has been measured during the signal strength database. Position $l$ which has the highest probability $P(S|l)$ for a signal strength vector is classified as the user's position in the field at this time.

$$P(S \mid l) = \prod_{i=1}^{|P|} \left( \frac{1}{\sqrt{2\pi\left(D_i^l\right)^2}} \exp\left( -\frac{\left(S_i - M_i^l\right)^2}{2\left(D_i^l\right)^2} \right) \right) \tag{4}$$

$$M_i^l = \frac{\sum_{i=1}^{n} RSS_i^l}{n} \tag{5}$$

$$D_i^l = \sqrt{\frac{\sum_{i=1}^{n} (RSS_i^l - M_i^l)}{n-1}} \tag{6}$$

### *2.3. Support Vector Machines (SVMs)*

Support Vector Machines are powerful techniques used for classification and data regression [12,13]. They are used for non-parametric supervised classifier for pattern recognition problems. SVMs are used in the localization system by training the support vectors on radio map that consist of grid points. SVMs analyse the relationship between the trained fingerprints and their grid points by considering each grid points as a class. The tested RSSI fingerprints are taken as an input to SVM that predict the class to which the tested belongs. This technique can be generalized to classify between more than two classes for $N$ training data $(x_i, y_i)$ .

Before any classification, the RSSI fingerprint vectors are mapped into higher dimensional space using kernel function. The SVM kernel function $K(..,..)$ is the dot product of two feature vectors $x_i$ and $x_j$ in some expanded feature space, there are several kernels are proposed by researchers. The four basics kernels as follow: linear, polynomial, sigmoid and radial basis function (RBF). In this research, linear is used in the following form

$$K(x_i, x_j) = x_i^T . x_j \tag{7}$$

Where $\sigma^2$ is the variance (i.e. width) of the Gaussian kernel.

After representing the training data by mapping the data to the feature space. The SVM algorithms identify hyperplane, which separates the support vector trained with a distance equal $\frac{2}{\|w\|}$. It is constructed in such a way that they can be divided in two data classes with a maximum distance to the closest vector from the same class. The optimization problem is shown in:

$$y_i (w^T x_i + b) - 1 \geq 0 \tag{8}$$

$$\max \frac{2}{\|w\|} \to \min \frac{\|w\|}{2} \to \min \frac{1}{2}\|w\|^2. \tag{9}$$

$$\min \left\{ L_{pd} (w, b, \alpha_i) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \alpha_i \left[ y_i \left( w^T x_i + b \right) - 1 \right] \right\} \tag{10}$$

Where $b$ and $\alpha_i$ are solution of the constrains and $y_i$ is the output of each class $y_i \in \{1, -1\}$, which achieve the minimize (9) based on lagrangian function, where $\alpha_i$ is the lagrangian multipliers. The constrained optimization problem can be expressed in a dual form by searching a solution under the form [38].

$$w = \sum_{i=1}^{N} \alpha_i y_i x_i \qquad \sum_{i=1}^{N} \alpha_i y_i = 0 \tag{11}$$

Maximizing with respect to $\alpha$ :

$$\max \left\{ \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i . x_j^T) \right\} \tag{12}$$

Under constrains:

$$\sum_{i=1}^{N} \alpha_i y_i = 0 \text{ Where } \alpha_{i \geq 0} \quad \forall i \tag{13}$$

There is a pure decision and uncertainty is inefficient. A pure decision means each node's data subset contains one and only one target location.

## 3. Results and discussion

The research material used indoor localization based Wi-Fi using Support Vector Machine method is the result of the measurement of signal strength received by laptop for IEEE 802.11g. The experiment was carried out in the corridor of the 3rd floor of the building with an area of ± 302 m2. Determination of reference points is the next stage in space planning which is the scope of research. At this stage the corridor is measured and then divided into areas with an area of 1 m$^2$. Before taking data training, the reference points that are right in the middle of each area are marked first and it is certain that the mark is in accordance with the coordinates that will be used as classifications in data fingerprint, this is done to facilitate the process of measuring data training.

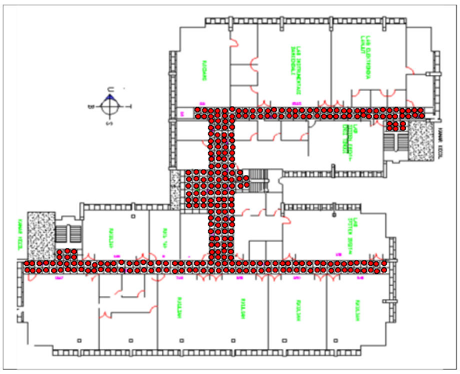


**Figure 1.** Reference points of 1 $m^2$.

This research is divided into several stages. These stages are:

- Space Planning is the first step in making a signal strength map in that space become the scope of research. The research room is floor hall 1, 2, and 3 a building. At this stage the aisle is measured then divided into cells with a width of 1 meter each cell.
- Measurement of RSS. In this process RSS measurements are received by the Laptop, at each cell that has been measured. Measurement done in the same direction i.e. west. Measurement RSS is done when the building conditions are quiet, that is on Saturdays and Sundays or holidays. Software the one used is Net Surveyor.

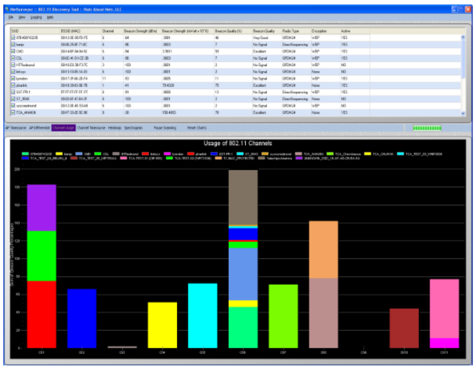


**Figure 2.** Netsurveyor view.

- RSS visualization. RSS visualization is used for provide a map of the signal strength received (RSS). RSS value on each grid point is obtained by calculating the average the signal received on the grid point. Visualization is done using software RapidMiner. Visualization using the AP installed on the 3rd floor of the building.

**Figure 3.** RSS visualization for AP1.

- Data Validation. Dividing the dataset for validation and for training, need to know whether the model to be made is the best model. Then finally use the statistical method to estimate the accuracy of the model against previously unknown data.

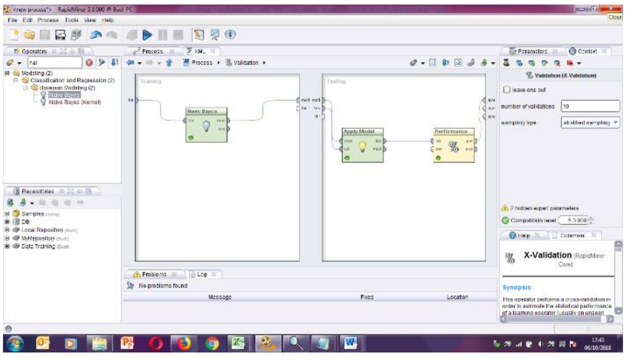


**Figure 4.** Data validation using RapidMiner.

- Location Fingerprinting. Fingerprinting is method for measurable data mapping, namely RSS to in a grid-point that covers the entire localization area which will be used for location estimation. RSS not all of the AP in the building environment used, but RSS is chosen from several APs has a significant influence on location estimation. RSS data processing is done using Excel software.
- Algorithm modelling. This process is carried out algorithm modelling used is Naïve Bayes and Support Vector Machine. In the process this modelling can be known the percentage accuracy.

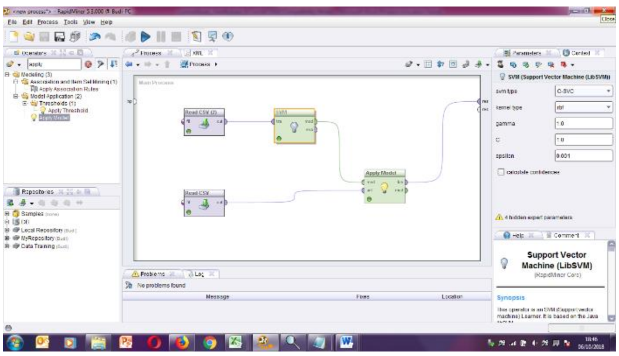


**Figure 5.** Algorithm modelling.

- Location Tests and Estimates. Testing is an important process for knowing results from a system. Test data obtained by how to do RSS measurements received laptop by walking along the 1st floor hallway, 2 and buildings. Estimated location of objects in this case is an IEEE 802.11 based laptop, obtained from a comparison between RSS measurements in fact, the test data with previous measurements that have been stored in fingerprint. Location estimation using algorithm which has been modelled before. Estimated error location is obtained by calculating distance between the actual location and the estimated location.

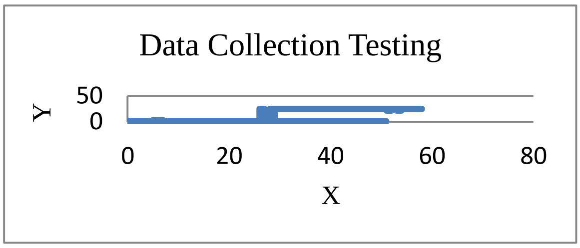



**Figure 6.** Data collection testing.

- Analysis of results. At this stage an analysis is carried out for knowing the magnitude of the estimated error on algorithm used.

Results of research in figure 7 shows the Class for indoor collection testing data for localization based Wi-Fi___33 signal using a support vector linear machine. Data testing was collected in the number of 10 to 50 data taken by walking along the hallway area in the building.

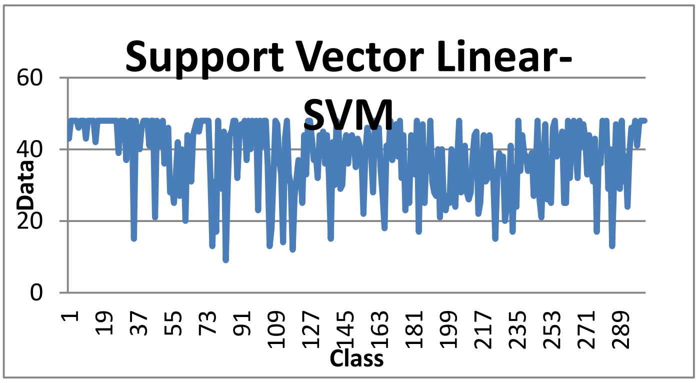



**Figure 7.** Class of testing data.

Table 1 shows the results of the minimum distance error by comparing the differences in the methods used namely Naïve Bayes and Linear SVM with C = 1, 2 and 5.

**Table 1.** Comparison of the minimum average distance with Naive Bayes and Linear SVM methods.

| Algorithm | Error (m) |
|---|---|
| Naïve Bayes | 6,21 |
| Linear SVM (C=5) | 5,884873 |
| Linear SVM (C=1) | 4,612462 |
| Linear SVM (C=2) | 4,612462 |

## 4. Conclusions

The SVM Linear method is better in the results of the minimum average error distance compared to Naïve Bayes. Linear SVM with C = 1, 2 and 5 shows that C = 1 and 2 get a fixed result but when C = 5 shows the change in the results of the minimum average error distance is not good compared to C = 1 and 2.

## References

[1] Hopper B A and Harter T 1993 The active badge system *InProc. Of INTERCHI-93* 533–534
[2] Kindberg F T 2002 System software for ubiquitous computing *IEEE Pervasive Comput.* **1** 1 26–35
[3] Bahl and Padmanabhan V N 2000 RADAR: An in-building RF-based user location and tracking system *INFOCOM* 775–784
[4] Elnahrawy R M E and Li X 2004 The limits of localization using signal strength: A comparative study *Proc. 1st IEEE Int. Conf. Sens. Ad Hoc Commun. Networks, St. Clara, CA*
[5] Ladd D W A, Bekris K, Marceau G, Rudys A and Kavraki L 2002 Robotics-based location sensing using wireless ethernet *Proc. Tenth ACM Int. Conf. Mob. Comput. Netw.*
[6] Aalto T O L, Gothlin N and Korhonen J 2004 Bluetooth and WAP push based location-aware mobile advertising system. In MobiSYS ’04 *Proc. 2nd Int. Conf. Mob. Syst. Appl. Serv.* 49–58
[7] Priyantha H B N B and Chakraborty A 2000 The cricket location-support system *Mob. Comput. Netw.* 32–43
[8] Ward A H A and Jones A 1997 A new location technique for the active office *IEEE Pers. Commun.* **4** 5 42–47
[9] Kupper A 2005 *Location Based Services: Fundamentals and Operations* (John Wiley & Sons)
[10] Caffery J 2000 *Wireless Location in CDMA Cellular Radio Systems* (Kluwer Academic Publishers)
[11] Kamol K and Prashant K 2004 Properties of Indoor Received Signal Strength for WLAN Location Fingerprinting *Proc. First Annu. Int. Conf. Mob. Ubiquitous Syst. Netw. Serv.*
[12] Schölkopf B and Smola A 2002 *A. Learning with Kernels* (MIT Press, Cambridge, MA)
[13] Joachims T 1999 Transductive inference for text classification using support vector machines *Proceeding Int. Conf. Mach. Learn.* 200–209
[14] Madigan A K D, Elnahrawy E, Martin R P, Ju W H and Krishnan P 2005 Bayesian indoor positioning systems *INFOCOM. IEEE* 1217–1227
[15] Ash J and Potter L 2004 Sensor network localization via received signal strength measurements with directional antennas *Conf. Commun. Control. Comput.* 1861–1870
[16] King W E T, Kopf S, Haenselmann T and Lubberger C 2006 Compass: A probabilistic indoor positioning system based on 802.11and digital compasses *WiNTECH* 34–40
[17] Youssef A U S M A and Agrawala A 2003 *Wlan location determination via clustering and probability distributions* (Washington, DC, USA)
[18] Elnahrawy R P M E and Austen-Francisco J 2007 Adding angle of arrival modality to basic RSS location management techniques *ISWPC* 464–469
[19] Niculescu D and Nath B 2004 VOR base stations for indoor 802.11 positioning *MobiCom* 58–69
[20] chung P, Chan C Y T, Tsui W Y and So H C 2006 Time-of-arrival based localization under NLOS conditions *IEEE Trans. Veh. Technol.* **55** 1