ACM Word Template for SIG Site

1st Author

1st author's affiliation  
1st line of address  
2nd line of address  
Telephone number, incl. country code

1st author's E-mail address

2nd Author

2nd author's affiliation  
1st line of address  
2nd line of address  
Telephone number, incl. country code

2nd E-mail

3rd Author

3rd author's affiliation  
1st line of address  
2nd line of address  
Telephone number, incl. country code

3rd E-mail

**ABSTRACT**

Attributed to the omnipresence of the radio signals for communications, sensing and recognition utilizing the Wi-Fi signals has significant advantage in terms of accessibility over conventional sensing means such as the camera. However, utilizing the raw Wi-Fi signals to capture in-air handwritten signatures for identity verification is yet a challenging task. In this paper, we propose a system for identity verification based on the handwritten signature signals captured by the Wi-Fi Channel State Information (CSI). A triplet network is adopted to learn the correlation between the captured signals and the user identities. To facilitate a fast converging loss model, a kernel and the range space learning is initially adopted for mining the triplet inputs. Subsequently, the triplet network is trained on a ConvNet structure based on the mined triplet inputs. Our experiments on a Wi-Fi dataset collected in-house show encouraging verification accuracy with faster training loss convergence comparing with that of the baseline triplet network and the Siamese network.

**CCS Concepts**

**• Information systems➝Database management system engines**   • **Computing methodologies➝Massively parallel and** high-performance simulations. **This is just an example, please** use the correct category and subject descriptors for your submission*.* The ACM Computing Classification Scheme:

<http://www.acm.org/about/class/class/2012>. Please read the [HOW TO CLASSIFY WORKS USING ACM'S COMPUTING CLASSIFICATION SYSTEM](http://www.acm.org/publications/article-templates/CCS-HOWTO-v6-12Jan2015.docx) for instructions on how to classify your document using the 2012 ACM Computing Classification System and insert the index terms into your Microsoft Word source file.

**Keywords**

Wi-Fi signature signal; in-air handwritten signature verification; the Kernel and the Range space projection learning; triplet network

# INTRODUCTION

Over recent years, several behavioral biometric traits have attracted attention in view of their rigid physical body independence. Among these behavioral biometrics, the signature-based user authentication [1, 2, 3] has attracted considerable interest with the development of in-air signature recognition systems [4, 5, 6, 7]. With the help of sensors such as the depth camera [4, 5] or a mobile sensor [6], the in-air signature recognition system has lower the spatial constraint in the process of signature acquisition comparing with contact-based authentication systems [8, 9, 10].

Recently, the commercial Wi-Fi device has been adopted for in-air signature authentication due to its easy accessible property [7]. Based on the distortion of the Wi-Fi CSI signal according to the user’s gestures, the in-air signature recognition system showed reasonable user verification performance [7]. More recently, some studies attempted to implement the deep learning algorithms in Wi-Fi signal-based user authentication systems to improve the verification performance [11, 12].

In this paper, we utilize a deep triplet network for identity verification based on the Wi-Fi CSI signature signal. To achieve not only the desired verification accuracy but also a fast training speed, we adopt the kernel and the range (KAR) space learning [13, 14, 15, 16] in order to mine the distinctive triplet inputs. Subsequently, the triplet network which utilizes the ConvNet [17] structure as a feature extractor is trained based on the L2 distance comparison.

The main contributions of our work can be summarized as follows:

* Proposal of a system for identity verification based on the Wi-Fi handwritten signature signals using a deep triplet network.
* Adopted the kernel and the range (KAR) space learning in order to mine the distinctive triplet inputs which boosted the convergence speed of the training loss in the triplet network.
* Provision of an experimental study using a Wi-Fi handwritten signature dataset which was collected in-house based on 50 subjects.

The paper is organized as follows: related works including the triplet network and KAR space learning are introduced in Section 2 for immediate reference. Our proposed method is discussed in Section 3. Section 4 describes our experimental results and analysis. Some concluding remarks are given in Section 5.

# Related works

## Triplet network

The triplet network is considered a metric learning based model which aims to learn useful representations by means of distance comparison [18]. It is often seen in person re-identification [19, 20, 21, 22] where the individual identities are matched based on discriminative image features. The main difference between person identification and re-identification is that the later is a more challenging task where images of the same person taken from different cameras or under different occasions are to be associated. In order to address our challenging Wi-Fi based verification task, we adopt a triplet network which optimizes the input data space so that data points with the same identity are closer to each other than those with different identities [23].

The triplet network receives triplet pairs of data as its input. These data triplets are constructed based on a combination of the input data. Since not all triplet samples contribute to the desired classification, recent attention has been paid to the choice of relevant input pairs for training. In order to optimize the training process which utilizes only some parts of the triplet pairs, several researches [20, 21, 22] generated triplets from a small number of classes (persons) in each iteration. In [24], a triplet mining process was implemented to speed up the training convergence. They utilized a large mini-batch at each training iteration and selected the triplets based on network training instead of random sampling. However, this strategy needed a few thousands of exemplar mini-batches in every training iteration for triplet pairs selection. This results in a heavy computational load in training. In order to make use of the small sample training size as well as to speed up the triplet training process, we adopt the kernel and the range space method for the learning.

## Kernel and the range space learning

Generally, the multilayer feedforward neural networks is trained based on the gradient descent method via backpropagation [25]. However, setting the learning parameters such as the learning rate and the learning momentum is a time consuming task.

Recently, a gradient-free learning framework based on the kernel and the range (KAR) space manipulation has been developed for multilayer network learning [13, 14, 15, 16]. The learning method is grounded on linear algebra without needing any training iteration.

Given training samples. Let denotes the training data set and denotes the network output. Then the multilayer neural network can be written as follows:

(1)

where ,, are the network weight matrices, is the bias vector, and is the activation function. By adopting an one-hot encoded target , training of the weight matrices using the KAR space method [16] can be computed as follows:

= 1,…,n. (2)

# Proposed System

In this section, we propose an identity verification system based on the Wi-Fi in-air handwritten signature (which will be called Wi-Fi signature hereafter) using the triplet network [18]. Fig.1 shows an overview of the proposed system utilizing the kernel and the range (KAR) space learning [14, 16] for mining the triplet inputs. Essentially, the KAR space projection learning is utilized to learn the triplet input data by mining those hard positive and the hard negative samples from each of the given anchor sample (see item (a) in Fig. 3). The hard positive and the hard negative samples refer to positive and negative class samples which are likely to be misclassified by the network. Subsequently, the ConvNet structure in the triplet network (see item (b) in Fig. 3) is trained with the mined triplet data based on a triplet loss function using the L2 distance comparison (see item (c) in Fig. 3). The following subsections describe the details of the triplet mining using KAR space learning and the triplet network.

## Triplet mining using the kernel and the range space learning

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

*ICIME 2017,* October 9–11, 2017, Barcelona, Spain

© 2017 Association for Computing Machinery.

ACM ISBN 978-1-4503-5337-3/17/10…$15.00

DOI: https://doi.org/10.1145/3149572.3149592

The network receives a triplet set of data as its inputs. These triplet data consist of the reference data (will be called anchor samples hereafter) and the corresponding positive class data (same class with that of the anchor) and the negative class data (different class from that of the anchor). The goal of the triplet network is to position the feature vectors with appropriate separation space by putting the positive samples close to the anchor sample while keeping the negative sample away from the anchor sample.

According to [24], it is important to select the hard positive samples and the hard negative samples with reference to the given anchor sample for fast loss convergence when training the triplet network. A hard positive sample is defined as a sample whose distance to the anchor sample is large (which is most likely to be misclassified as a negative sample). On the other hand, a hard negative sample is defined as a sample whose distance to the anchor sample is small (which is most likely to be misclassified as a positive sample). However, from the raw data, there is no information regarding whether a sample is considered hard positive or hard negative before we train the network.

In this work, we propose to adopt the kernel and the range (KAR) space learning (see Section 2.2 for details) as a pretraining network to mine the hard positive/negative samples from the given anchor sample. Since the KAR space learning has no iterative learning process, we can mine the triplet samples without using the time consuming backpropagation training process.

By training the network with the single shot KAR space learning, we can map the distance among the samples by using the output vector of the KAR space network. Given a set of training data which is packed in matrix , the network output can be written as:

(3)

After training based on the KAR space projection, the given anchor sample can be used as the reference to decide whether the network output of a sample is far away from this anchor sample. In other words, a hard positive sample and a hard negative sample with respect to the anchor sample can each be determined based on:

(4)

(5)

where and respectively denote the thresholds that determine whether a sample is hard positive or hard negative. Since the hardest samples are likely to be outliers which can degrade the training process of the triplet network, we empirically set the at 75 percentile of the distance and at 25 percentile of the distance. The final set of and samples is randomly selected based on equations (4) and (5).

## The ConvNet structure

The next step is to design a feature extractor which converts the triplet input data into feature vectors. In this work, we utilize the ConvNet structure [17] as a feature extractor since the three dimensional data format of our preprocessed input signal can be regarded as an image data format with multiple channels.

Our ConvNet design (item (b) in Fig 3) consists of three convolutional layers and one fully-connected layer. The number of convolutional filters to be trained in each layer is empirically chosen as , with fixed filter size of , each of stride 1. The Rectified Linear (ReLU) activation function and the Max-pooling layer are applied between each convolutional layer. Subsequently, the extracted features from the last convolutional layer are flattened into a vector before feeding into the fully-connected network. The output vectors from the fully-connected layer are finally transformed using the sigmoid function followed by a normalization.

## The triplet loss

The triplet loss function was first seen in [18] for training the triplet network. For the network input , an anchor sample is randomly selected among the training data set. The positive and negative samples (respectively and ) are then determined and being selected for training the ConvNet based on the feature vectors , for i=1,...,N. The triplet loss function is formulated based on a summation of the difference between the positive distance (the distance between the anchor vector and the positive vector) and the negative distance (the distance between the anchor vector and the negative vector) as follows:

(5)

## Subsequent Pages

For pages other than the first page, start at the top of the page, and continue in double-column format. The two columns on the last page should be as close to equal length as possible.

Table . Table captions should be placed above the table

|  |  |  |  |
| --- | --- | --- | --- |
| **Graphics** | **Top** | **In-between** | **Bottom** |
| Tables | End | Last | First |
| Figures | Good | Similar | Very well |

## References and Citations

Footnotes should be Times New Roman 9-point, and justified to the full width of the column.

Use the “ACM Reference format” for references – that is, a numbered list at the end of the article, ordered alphabetically and formatted accordingly. See examples of some typical reference types, in the new “ACM Reference format”, at the end of this document. Within this template, use the style named *references* for the text. Acceptable abbreviations, for journal names, can be found here: <http://library.caltech.edu/reference/abbreviations/>. Word may try to automatically ‘underline’ hotlinks in your references, the correct style is NO underlining.

The references are also in 9 pt., but that section (see Section 7) is ragged right. References should be published materials accessible to the public. Internal technical reports may be cited only if they are easily accessible (i.e. you can give the address to obtain the report within your citation) and may be obtained by any reader. Proprietary information may not be cited. Private communications should be acknowledged, not referenced (e.g., “[Robertson, personal communication]”).

## Page Numbering, Headers and Footers

Do not include headers, footers or page numbers in your submission. These will be added when the publications are assembled.

# Experiments

## Dataset

In order to evaluate the verification performance of the proposed system, the Wi-Fi CSI signature dataset [7] is utilized in our experiments. The Wi-Fi CSI signature dataset consists of 2000 Wi-Fi CSI signature signals (4 directions 10 samples 50 identities) with sample size 500306. We utilize only the absolute value from each complex CSI signal in our experiments.

Captions should be Times New Roman 9-point bold. They should be numbered (e.g., “Table 1” or “Figure 2”), please note that the word for Table and Figure are spelled out. Figure’s captions should be centered beneath the image or picture, and Table captions should be centered above the table body.

## Experimental settings

# SECTIONS

The heading of a section should be in Times New Roman 12-point bold in all-capitals flush left with an additional 6-points of white space above the section head. Sections and subsequent sub- sections should be numbered and flush left. For a section head and a subsection head together (such as Section 3 and subsection 3.1), use no additional space above the subsection head.



Figure 1. Insert caption to place caption below figure.

.

## Subsections

The heading of subsections should be in Times New Roman 12-point bold with only the initial letters capitalized. (Note: For subsections and subsubsections, a word like *the* or *a* is not capitalized unless it is the first word of the header.)

### Subsubsections

The heading for subsubsections should be in Times New Roman 11-point italic with initial letters capitalized and 6-points of white space above the subsubsection head.

#### Subsubsections

The heading for subsubsections should be in Times New Roman 11-point italic with initial letters capitalized.

#### Subsubsections

The heading for subsubsections should be in Times New Roman 11-point italic with initial letters capitalized.

# ACKNOWLEDGMENTS

Our thanks to ACM SIGCHI for allowing us to modify templates they had developed.

# REFERENCES

1. Bowman, M., Debray, S. K., and Peterson, L. L. 1993. Reasoning about naming systems. *ACM Trans. Program. Lang. Syst.* 15, 5 (Nov. 1993), 795-825. DOI= <http://doi.acm.org/10.1145/161468.16147>.
2. Ding, W. and Marchionini, G. 1997. *A Study on Video Browsing Strategies*. Technical Report. University of Maryland at College Park.
3. Fröhlich, B. and Plate, J. 2000. The cubic mouse: a new device for three-dimensional input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (The Hague, The Netherlands, April 01 - 06, 2000). CHI '00. ACM, New York, NY, 526-531. DOI= <http://doi.acm.org/10.1145/332040.332491>.
4. Tavel, P. 2007. *Modeling and Simulation Design*. AK Peters Ltd., Natick, MA.
5. Sannella, M. J. 1994. *Constraint Satisfaction and Debugging for Interactive User Interfaces*. Doctoral Thesis. UMI Order Number: UMI Order No. GAX95-09398., University of Washington.
6. Forman, G. 2003. An extensive empirical study of feature selection metrics for text classification. *J. Mach. Learn. Res.* 3 (Mar. 2003), 1289-1305.
7. Brown, L. D., Hua, H., and Gao, C. 2003. A widget framework for augmented interaction in SCAPE. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology* (Vancouver, Canada, November 02 - 05, 2003). UIST '03. ACM, New York, NY, 1-10. DOI= <http://doi.acm.org/10.1145/964696.964697>.
8. Yu, Y. T. and Lau, M. F. 2006. A comparison of MC/DC, MUMCUT and several other coverage criteria for logical decisions. *J. Syst. Softw.* 79, 5 (May. 2006), 577-590. DOI= <http://dx.doi.org/10.1016/j.jss.2005.05.030>.
9. Spector, A. Z. 1989. Achieving application requirements. In *Distributed Systems*, S. Mullender, Ed. ACM Press Frontier Series. ACM, New York, NY, 19-33. DOI= <http://doi.acm.org/10.1145/90417.90738>.

Columns on Last Page Should Be Made As Close As Possible to Equal Length

Authors’ background

|  |  |  |  |
| --- | --- | --- | --- |
| Your Name | Title\* | Research Field | Personal website |
|  |  |  |  |
|  |  |  |  |

\*This form helps us to understand your paper better, the form itself will not be published.

\*Title can be chosen from: master student, Phd candidate, assistant professor, lecture, senior lecture, associate professor, full professor