Identifying Interaction Groups Using The Bluetooth Proximity Data Of The Conflab Dataset

Carlotta Lichtenauer
Supervisor(s): Stephanie Tan, Jose Vargas Quiros
Responsible professor: Hayley Hung
EEMCS, Delft University of Technology, The Netherlands
22-06-2022

A Dissertation Submitted to EEMCS faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering

Abstract

Detecting social interactions through wireless wearable Bluetooth devices is increasing in popularity. Devices use the signal strength to other detected devices to estimate the proximity between people and group them together based on the Dominant set algorithm. Dominant sets are a maximal clique of nodes with an edge-weight based on the affinity between the nodes. Nevertheless, the signal is heavily influenced by external factors, which increase in an crowded environment. This paper introduces three different noise reduction filters that try to detect the kind of noise and therefore improve the detection of surrounded devices. Further, this paper looks at the overall impact of proximity on the resulting RSSI values. Knowing this relationship helps to normalize the values and therefore eliminates the need to apply noise reduction. Using a dataset of 48 sensors recorded in a conference setting with a specific designed sensor the low frequency pass filter gets an accuracy score of 81.8% with a cut-off frequency of 0.07 Hz. It performs best when considering a conversation window of 20 seconds. Here, only 2/3 of the detected groups has to coincide with the actual formed group at a specific timestamp. Furthermore, the orientation of the participants to each other has heavy influence on the resulting RSSI values and therefore a normalization based on only proximity cannot be done.

1 Introduction

Understanding social behaviour and interactions contribute to learning about human society and consequently adds to better cooperation between people in many disciplines [1]. One way to detect these behaviours is by looking at nonverbal interactions [2]. This project assumes that participants are having an interaction when they actively engage with another participant over a certain amount of time. Nevertheless it is difficult to study the actions of individuals when the subjects know they are being recorded [2], [3]. Further, self-recorded data can easily be influenced by bias and different perceptions of the subjects [1]. Gedik [3, p 9] argues to overcome this problem the need for a sensor that is integrated into people's life stands.

For this dataset a sensor called "Midge "was built by the Socially Perceptive Computing Lab of the TU Delft [4]. It is a sensor which monitors human behaviour by collecting multimodal data. The Conflab dataset is the dataset used in this project. It was recorded with the participants wearing the sensor around their chest area. The setup for the data set was a conference with around 50 participants. The goal was for people to form several F-Formations to imitate a real conference scenario. F-Formations are described by Hung et al. as "a specific instance of a group of people who are congregated together with the intent of conversing and exchanging information with each other "[5, p 1]. To better visualize this concept 5 F-Formations were circled in red in Figure 1.

Figure 1: A snapshot taken of the conference experiment. 5 different F-Formations are denoted with a red circle.

These F-Formations are formed based on the proximity of the participants. Looking at human proximity is a major contributor when understanding a social setting [6] and the relationship between people [7]. This can be done by detecting surrounded Bluetooth devices. The recorded RSSI (Received Signal Strength Indicator) values from the Bluetooth signals allow researchers to make a distinction of possible face-to-face interaction based on the proximity and do not require to know the absolute position of the subject [8].

Gedik describes that distinguishing interactions becomes especially difficult when considering real-life situations, such as, crowded places [3, p 30]. This difficulty is due to the challenge in understanding with whom a person is having an interaction due to external factors such as no face-to-face interchange, other people standing close by, etc. Since this data set is recorded in a crowded setting, this project especially focuses on finding F-Formations and not face-to-face interactions.

Nevertheless, external factors can introduce a lot of noise in the RSSI data. This noise can be added through surrounding walls [9], the way people stand to each other [10] or just the fact that the sensor might be blocked by folded arms, jackets, moving hands, etc. These external factors might alter recorded data or at least have an impact on it and therefore it is crucial to understand their relation to the recorded data.

To overcome this problem, this project used three different noise reduction filters to preprocess the data before the grouping algorithm. A throughout analysis was done to learn how different parameter values of the applied filters affect the results. Further, the concept of a interaction window is introduced which assumes that people are in an interaction for a certain amount of time. Lastly, this research checked the actual impact of proximity on the resulting RSSI values and came to a conclusion if only proximity values can be used to find F-Formations.

F-Formations were detected by only considering the RSSI values recorded during the experiment. Interaction groups are identified by looking at the distances between participants based on the RSSI values. Based on a threshold, participants were considered to form a possible interaction group. The final decision, if two sensors were in the same group, was based on the resulting affinity matrix per timestamp and a group detection algorithm which found the dominant sets in the data. Dominant sets are maximal cliques in a weighted graph [11]. This weighted graph consists of nodes which represent the participants and edges which have a certain weight based on their affinity [12]. Participants are grouped by considering a certain threshold on the edge weights.

The main contributions of this paper are:

- Comparison between different noise reduction algorithms and their influence on the grouping result.
- A new approach to detect possible RSSI values which could contribute to two sensors being in an F-Formation.
- Different threshold approaches to detect an interaction group.
- Relationship between the distance of participants and the corresponding RSSI value of the data set.

2 Related Work

Bluetooth RSSI values have been used in different amount of settings to detect proximity between humans [2], [7], [9], [10]. It especially became a relevant approach to estimating the distance between people during the Coronavirus pandemic. Leith et al. looked into how distance and rotation between people affect the RSSI values [10] . Here the problem occurred that the RSSI values can not detect if people were close to each other or were for example separated by walls. Their results show that RSSI values increase in an indoor setting, this is most likely to the reflection of the signal on walls, etc. and people standing closer to each other. But consequently, the values are noisier. Relative orientation of the sensors plays another important role on the perceived RSSI value. For example, if the sensor is placed in on the back in a bagpack of the subject or in the pocket of their pants which is closer to the other sensor, has an influence of the strength of the value [8].

Nevertheless, the earlier mentioned experiments focused more on determining the relative distance between subjects rather than analyzing if participants might be interacting based on their distance to each other. Koshiba et al. tried to analyze people's behaviour during lunch based on the distance between the subjects and therefore contribute to more interaction between co-workers [7]. The MatchNMingle data set looked at social behaviour at a speed dating event with a cocktail party afterwards. Their research tried to determine how the one to one conversations during the dates could have had an influence on the proximity between subjects during the crowded cocktail party. In contrast to the research conducted in this project, [2] uses all the multimodal data retrieved from the sensor to detect social behaviour. Their results imply that adding the proximity factor to detecting face-to-face interactions increased the accuracy in detecting conversation groups and therefore supports the hypothesis that RSSI values can detect F-Formations [2].

RSSI signal values are easily influenced by external factors which lead to faulty results. To reduce the noise introduced by these factors noise filtering or smoothing should be applied. Dong et al. used three different techniques to smooth the signal; a moving average method, a weighted average method and a curve fitting method. However, the paper concluded that all smoothing techniques did not contribute much to detecting proximity between subjects and stated that RSSI should not be used as the only input when wanting to detect distance in an indoor setting [13].
Other papers do show the need of applying noise filters. Kosihba et al. used an averaging filter which filtered the values every minute to filter out by passing devices. Their results show a clear connection between physical proximity and the resulting RSSI values [7]. Instead of applying an averaging filter Mizarei et al. attempted to reduce noise by applying a median filter of length 15 and replacing all unknown values with a value of -100 dBm. They introduce a term called multipath fading which refers to the fact that external noise is added to RSSI values when considering signals [14]. Jianyong et al. argues that RSSI values have a lot of fluctuations due to reflection and therefore seem random. After multiple experiments, they indicated that the probability of another sensor being a certain distance away has resulting RSSI values which look like a Gaussian distribution [15]. Therefore, it makes sense to use a Gaussian filter to filter the randomness out of the signal.

Despite applying the noise filters, the relation between RSSI values and the general distance between interaction partners is helpful to know when grouping participants. Research by Liu et al. shows a clear connection between the signal strength and the distance between the sensors, which is visualized in Figure 2 [8]. Different papers have approached a distance equation which takes the signal values into account [8], [13], [16]. Nevertheless, they emphasize that noise still has a big impact on the distance value.

Furthermore, Liu et al. explains that it is important to consider continuous values since just considering just a simple threshold introduces a lot of error. Their final result shows a threshold of -52 dBm of an average distance of 1.52 cm [8]. Thus, it must be considered that they recorded these values in simple conversation settings and not in a crowded environment.

3 Methodology

3.1 Dataset

This project was conducted using the proximity data recorded with the Midge developed by the Social Perceptual Lab of the TU Delft [4]. The Midge is a specific developed sensor which records multimodal data. It is developed for a subject to wear around the chest area. Every Midge recorded the RSSI values to all other midges used in the experiment with the timestamp when the data was recorded. This project works with the recorded data of the Midge. Additionally, the true grouping data was provided by the Lab to match the groups found with the Dominant set algorithm to the actual formed groups.

Figure 3 shows that the ground truth data and the data recorded at a certain timestamp do not coincide. This can

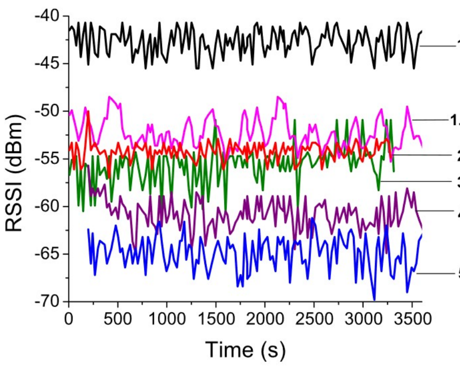

Figure 2: Distance measurements based on RSSI signals. Reprinted from [8].

have multiple reasons. When too many midges are recorded probably midges were not used in the experiment but still turned on. On the other hand, Midges could have fallen out during the experiment and therefore did not record any data. To fix the problem of too many recorded Midges it was checked if Midges were actually included in the ground truth. A solution for the problem that Midges did not record data was found in two parts. First, the data was interpolated by using a built in interpolation method of a 1D signal by scipy [17], which let for filling in unknown values, by drawing a line between two known data points. Secondly, a median filter is introduced in subsection 3.2 which replaces every value data certain timestamp with the median of certain kernel size.

Interpolation of the data also helped aligning the recorded data with the ground truth. The ground truth was recorded every second. Yet, the data set was recorded at random times with a maximal time span between recorded values of 4 seconds. The interpolation helped to get missing results at certain timestamps, alternatively making it easier to evaluate the data.

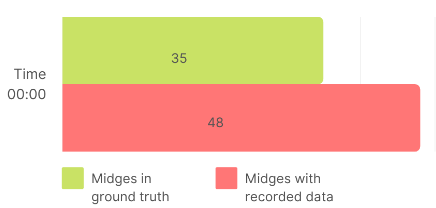

Figure 3: Bar plot of the Midges considered in ground truth vs how many Midges give data at a specific timestamp.

3.2 Noise reduction

This project evaluated three different noise reduction or noise smoothing filters. All filters were evaluated on how they perform considering different kernel sizes.

Median filter The first filter got introduced by [14]. Different papers used similar approaches for noise reduction with mean filters, averaging filters [6], [13]. This project chose to use the median filter since participants become interaction partners if they are at a certain distance from a sensor for a certain amount of time. Consequently, the median during this time will not deviate much from the expected values. Furthermore, will the median filter solve the problem that sometimes sensors were turning off and on and resulting in missing data.

Gaussian filter The second filter is a Gaussian smoothing filter. This filter is inspired by the experiment in [15]. Jianyong et al. argues that the probability of another sensor being a certain distance away has resulting RSSI values which is similar to a Gaussian distribution [15]. Therefore, it makes sense to smooth these signals with a Gaussian filter.

Low pass filter The last filter is a standard noise reduction filter. It has not been introduced yet by other research papers but it makes sense to assume based on the definition of an interaction, that the signal of two people which are in the same F-Formation should not vary too much. Furthermore, the Gaussian filter, which was introduced earlier, is a low frequency filter but with given kernel values. Jianyong et al. argued that RSSI signals are random. It is to assume that this randomness, which can be seen as noise, gets introduced when two sensors are far away, due to reflection, people walking through the signal, etc. [15]. Therefore, it makes sense to design a different low pass filter with different cutoff frequencies to determine how much the signal fluctuates. The filter uses a sample frequency of 1 Hz and uses different cut-off frequencies of 0.02, 0.07, 0.1, 0.3.

Relationship between distance and RSSI values Lastly, this project looks into the fact that the RSSI values might not be noisy at all, but rather tries to understand special values in the signal. This is based on the idea that there is a relation between the RSSI values and the distance between interaction partners, as shown in Figure 2. Looking at this relation it can be found out how much the Bluetooth RSSI values actually contribute to the findings of correct F-Formations. Here certain snippets in the videos will be compared to the ground truth with an analysis of the distance, orientation and corresponding RSSI value between participants.

3.3 Interaction window

To evaluate meaningful RSSI values, a threshold was introduced over a certain time window. This approach was chosen to introduce the idea that participants were expected to interact with each other over a given time range before they were grouped together. The experiments evaluate how different time windows influence the results with a threshold of -55 dBm. This threshold was chosen by conducting different research papers. Liu et al. clearly indicates in Figure 2, that when considering people with a distance of 1 - 2 meters away from each other it makes sense to consider the RSSI values ranging between -40 dBm and -55 dBM [8]. The final decision was based on the threshold identified by Dikker, another researcher who has worked with the Conflab dataset [18].

$$A[i][j] = \begin{cases} 1, & \text{if } \overline{x} \geq -55dBm \\ 0, & \text{otherwise} \end{cases}$$

Participants were grouped together if the mean value in a certain time window would be over the mentioned threshold. The mean value was chosen since it was assumed that even with fluctuations in the signal the overall mean during an interaction will be higher than the chosen threshold. The resulting discrete value gets filled into an affinity matrix, here denoted as A, for each timestamp, indicating if two sensors are in an F-Formation or not. The affinity matrix is an n x n matrix where n is the amount of sensors used in the experiment.

3.4 Dominant Sets

The grouping algorithm of the F-Formations is based on a clustering algorithm introduced by Pelillo et al. and used by Swofford et al. and Hung et al. [5], [11], [12]. Pelillo et al. describes a cluster or clique as group where the nodes have high homogeneity and high inhomogeneity to the nodes in other clusters [11, p 167]. The algorithm introduces a node for every participant and creates edges between these nodes which represent their neighbouring relationship. Every node gets a weight based on their affinity mentioned in subsection 3.3. According to a clustering algorithm, which groups node together where the overall similarity is higher than to the external nodes, dominant sets are formed. Dominant sets are than formulated as an interaction group.

3.5 Evaluation

To evaluate the methods mentioned and different input values for certain variables the project looks at the precision and recall score compared to the ground truth data. Precision indicates how many True-positives and False-positives are getting found. Recall indicates how many True-positives and False-negative are found. Distinguishing between the two scores is important because it can be said at the end if the methods put participants wrongly in groups or if they discard wrongly participants of a group. Finally, the F1 score is calculated based on these two values to give an indication of how well the methods and values perform. The scores are calculated considering two different thresholds. First, the accuracy is based on how many complete correct groups are detected. A correct detect group will result in a True-Positive if the all participants in the group coincide with the ground truth. Further in this paper this will be considered a 1-to-1 comparison. The second experiments will calculate the accuracy if already 2/3 of the found sensors in a group match with the ground truth, further in this paper this will be referred to as a 1-to-2/3 comparison.

4 Experiments

To analyze the methods introduced in subsection 3.2 and subsection 3.3, three experiments were conducted.

Experiment 1 - Noise reduction The first experiment looked at the effect of noise reduction or signal smoothing. This is because the RSSI signal can be strongly influenced by external factors. To estimate the amount of noise and its impact on the results three different filters were introduced. The experiments were conducted with kernel sizes of 5, 10, 20 and 30 analysing the impact of the level of smoothness over the signal.

Experiment 2 - Interaction window The second experiment analyzed the impact of the sliding interaction window. The window had differing sizes from 20, 80, 120, 180 and 240 seconds. The idea of this experiment is to discard pairwise nodes which only walk past each other and do not interact with each other.

Experiment 3 - Relationship between distance and RSSI It was assumed that peaks and specific values in the signal get discarded when applying a noise filter. Therefore this project looked at specific sensors at certain timestamps and tries to find a relationship between the distance of participants and their resulting RSSI Values.

5 Results

The results of the three experiments mentioned in section 4 are shown in the following tables.

5.1 Results for the experiments

Experiment 1 and Experiment 2 The first and second experiments were conducted together since they both perform different parameter analysis and were likely to influence each other. The values are combined in a table which can be found in Appendix A in Table 6. A summary of the best performances for each noise filter can be found in Table 1. Figure 4 shows all three filters applied on the signal.

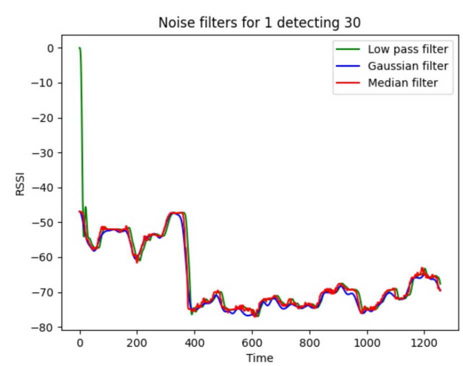

Figure 4: All 3 noise filters applied on the resulting RSSI signal from sensor 1 detecting sensor 30.

Precision	Recall	F1-score	Window	Kernel	Filter
0.92	0.6499	0.763	20 sec	∼	No
0.91	0.69	0.788	80 sec	10	Med
0.9269	0.717	0.808	20 sec	30	Gau
Precision	**Recall**	**F1-score**	**Window**	**Cut-off**	**Filter**
0.920	0.737	**0.818**	20 sec	0.07	Low

Table 1: Summary of the best results for every noise filter when conducting experiment 1 and 2 together when considering 2/3 of a detected interaction group to the ground truth.

When comparing 2/3 of an individual group to the ground truth, results verify the assumption that interaction people will have an interaction window of a certain time. The table shows that this window is especially useful when considering a size of 20 and 80 seconds. Moreover, the use noise filtering when having an interaction window is effective. The F1 score of the non filtered signal is informative but the scores improves when using noise filters. The best score of *0.818* is perceived with a *20* seconds window using a low pass filter with a cut-off frequency of *0.07* Hz.

Precision	Recall	F1-score	Window	Kernel	Filter
0.718	0.493	0.5854	20 sec	0	No
0.850	0.513	0.602	20 sec	20	Med
0.7361	0.555	0.6333	20 sec	20	Gau
Precision	**Recall**	**F1-score**	**Window**	**Kernel**	**Filter**
0.72580	0.575	**0.6417**	20 sec	0.07	Low

Table 2: Summary of the best results for every noise filter when conducting experiment 1 and 2 together when considering the whole detected interaction group to the ground truth.

Table 2 shows the results when considering the 1-to-1 comparisons of every found group to the ground truth, it can be seen that the F1 score is overall worse than the 2/3 comparison. Further, it is interesting to see that the overall time windows also decrease. The best value is *0.6417* with a low pass filter with a cut-off frequency of *0.07* Hz. These are the same parameter values as for the comparison.

Experiment 1 This experiment is established based on the results from conducting experiments 1 and 2 together. Since the window size is very small it is interesting to examine what influence the window has in general on the scores. The results in Table 3 compared to the results in Table 1 indicate that the interaction window together with noise reduction filters gives the best values. When considering no interaction window, having no filter gives the best result.

Precision	Recall	F1-score	Kernel size	Filter
0.93	0.639	**0.758**	~	No
0.928	0.612	0.738	20	Median
0.917	0.6352	0.75	30	Gaussian
0.919	0.618	0.739	~	Low

Table 3: Summary of the results without considering a sliding time window. These results show how much influence the concept of the sliding window on the accuracy results when detecting correct interaction groups. The parameter values are chosen based on the earlier conducted experiments.

Experiment 3 To understand the relationship between proximity and the resulting RSSI values, 2 situations were conducted: Firstly, the distance, orientation and RSSI values between subjects in a group and close by groups in a crowded environment. Secondly, how these values behave when two participants move next to each other over a time range.

First scenario Table 4 shows the orientation, distance and corresponding RSSI values of sensor 33 to all other participants seen in Figure 5. Sensor 33 is the person in the middle of the image. This person was chosen since he stands close to many other participants but only in an F-Formation with one other participant. Two value pairs are interesting to highlight in Figure 5. The numbers marked in yellow have similar orientations, but very different distances to sensor 33. Still, the resulting RSSI value is the same.

Furthermore, the values marked in green are 80 cm apart, however, the latter value is lower due to the orientation of the two people.

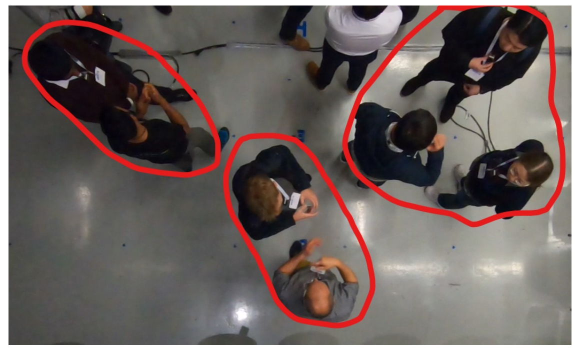

Figure 5: Group constellation for 3 F-Formations marked with red circles. Sensor 33 is the person in the middle of the image with a dark jacket, grey shirt and a white badge.

Sensor	Sensor	RSSI	Group	Distance	Orientation
33	4	-65	Y	20 cm	120 °
33	9	-	N	110cm	120 °
33	26	-65	N	90 cm	S-to-S
33	27	-68	N	40 cm	F-to-B
33	7	-70	N	120cm	F-to-B
33	34	-73	N	40cm	B-to-B

Table 4: Summary of values for scenario in Figure 5. These values are conducted by considering sensor 33 which is the person in the middle of the image. A row shows the relative position of sensor 33 to its surrounding sensors based on the physical distance and orientation.

Second scenario Figure 6 shows 4 constellations of two participants in the time range of 10 seconds. Table 5 summarises how far the two subjects stand away from each other, how they are rotated to each other and the resulting RSSI value. It can be seen that the rotation of the people has an impact on the RSSI value. Regardless, the RSSI value at the tenth second, marked in red, stands out compared to the results mentioned earlier. Because the people are standing back to face to each other and not in the same group but still results in a value of -43.

(a)
Constellation at
8 seconds

(b)
Constellation at
10 seconds

(c)
Constellation at
14 seconds

(d)
Constellation at
18 seconds

Figure 6: Interaction scenario in a time range of 10 seconds

Time	Sensor	Sensor	RSSI	Group	Distance	Orientation
0:08	35	19	-51	Y	20 cm	F-to-F
0:10	35	19	-43	Y	25 cm	B-to-F
0:14	35	19	-71	N	15 cm	S-to-F
0:18	35	19	-51	Y	10 cm	B-to-F

Table 5: 4 Scenarios of the relationship between sensor 35 and sensor 19. A row shows the relative position of the sensor 35 and 19 to each other based on their distance and rotation.

6 Discussion

This paper presented four different methods and contributions to analysing if the Bluetooth proximity data is a valid input to estimate F-Formations. The application of noise filters on the signal was inspired by multiple papers which used noise filtering on their experiment data. Noise filters reduce fluctuating signal values to be more stable making it possible to miss potential peak moments which were found when applying no sliding window and no noise filter. However, these constant values are detected by an interaction window which will take all values in the window range into account. That is why overall the F1 score was better when having a noise filter and a sliding interaction window combined. Overall the low pass frequency filter performed the best. The best F1-score of *0.818* gets conducted by applying a low frequency pass filter with a cut-off frequency of *0.07* Hz and a window size of *20* seconds. That supports the hypothesis on how the use of low pass filter is important for detecting participants which interact over a certain amount of time. Since the resulting signal does not fluctuate as much as when they are not in the same group. The reason why the cut-off frequency of 0.07 Hz performed the best and not the one of 0.02 Hz could be because the participants are still moving a bit while interacting. Consequently, the small movements are not discarded by this cut-off frequency.

Nevertheless, noise filters only show an improvement when used in combination with an interaction window. When not using this window the score value dropped to *0.739*. Overall smaller interaction windows gave the best results. This indicates that participants are not too long in the same F-Formation together. In the videos recorded during the data collection it can be seen that people tend to rotate their bodies during an interaction. Therefore, the window might only detect snippets of the interaction.

Furthermore, allowing a group to result in a True-Positive when *2/3* of the detected group coincides with the ground truth instead of having the whole group match gives always a better result. Here it is important to mention that the precision score is always higher than the recall score, meaning that there are more False-Negatives than False-Positives in the detected groups. This might be in connection to the sliding window used which discarded group members if they do not reach a certain threshold over an amount of time. Participants tend to move their body orientation a lot while interacting. Results have shown that the orientation has an influence on the resulting RSSI values and therefore subjects might not get detected.

When analysing the relationship between the distance and the corresponding RSSI value it got explicit that the orientation of two sensors to each other has a big impact on the resulting RSSI value. This got evident in the first scenario when looking at the values marked in yellow in Table 4. Two people who were similarly oriented to another sensor but had different distances resulted in the same RSSI value. These results coincide with the research conducted in [10]. The reason why RSSI values are influenced so much by the orientation can be caused by the signal being blocked when the sensors are not exactly facing each other.

On the other hand, the second scenario showed that the distance between participants can still have a big impact on the RSSI value. The lowest value can be found when the participants are facing each other back to face. This result clashes with the evaluation made that the orientation has a big influence on the resulting RSSI value. Another logical explanation for the value marked in red in Table 5 could be that the values are recorded with a delay and the value refers to 1 second before this movement happened.

Furthermore, it is important to mention that the relationship between proximity and RSSI, as well as the added noise heavily depends on the surrounding in which the data set was recorded. The Conflab set was in a crowded setting and therefore had consequently different values than values recorded in seated or walking settings.

Ultimately, the results support the need for an equation which takes the orientation, as well as the distance of subjects into account. Having established such an equation it would be straightforward how to normalize the RSSI values so that there is no need for a noise reduction filter but rather having a more precise cut-off threshold.

7 Conclusion and Recommendation

Using the proximity data based on the recorded RSSI values is a valid first step in detecting F-Formations in a mingling data set. Here, it must be taken into account that the resulting values can contain noise which is introduced through external factors, such as walls, the orientation of the person, other participants, etc. Therefore, it is crucial to add a noise filter. Since participants are expected to stay closer to each other while having an interaction a low frequency pass filter is suitable. To obtain better results, applying an interaction window of around 20 to 80 seconds in which the signal values must be above a determined threshold adds to the result.

Nevertheless, proximity is not the only value which influences the corresponding RSSI values. Especially in crowded environments can the orientation of interaction partners change constantly. This constant change had a big influence on the orientation of the person and is the reason why the interaction window is useful but only in small time ranges. In future research, it is recommended to combine the orientation of a subject with its proximity to be able to distinguish whether participants are in a group or not.

This research has opened the door for future work on this data set and the overall goal of analysing social behaviour based on multi-modal data in a crowded environment. It is noted that considering RSSI values alone already perform quite well on detecting correct interaction groups. Furthermore, this research has studied different noise filters and their effect on the accuracy of the grouping algorithm, as well as come to a conclusion why noise filters work well with the newly introduced concept of an interaction window. A noticeable result is the effect of the orientation on the RSSI value. This knowledge can be used in future research and experiments to combine the orientation with the proximity of interaction partners and consequently resulting in higher accuracy when detecting F-Formations.

8 Responsible Research

This paper has been produced while conducting scientific integrity. Hereby, it was especially looked at how ethical the data collection and data processing was. And how reproducible the data set and experiments conducted in this paper are.

Ethical behaviour To respect the privacy and confidentiality of the participants in the experiment, no personal data was recorded and every participant got an id number. Further, to match the sensors to the participants in the recorded video a description of their clothing was conducted. Based on the recorded data and the recorded information through the videos and experiment, it is not possible to trace personal information of the participants back.
Throughout the research, only the sensor id's have been used to identify the sensors and neighbouring sensors.

Reproducible research A major part of scientific integrity is to analyse how the experiments and results conducted in this research can be reproduced and therefore adapted to different scenarios, further experiments or a different data set. Having this reproducibility makes it easier to analyse how valuable the results found in this research are and put them in perspective to future results.

The details of how the experiments are conducted, which values were used and which specific methods were applied are in detail explained in this report. The result tables show a clear indication of which values were used to tweak the noise filters.

The data was recorded with a sensor specifically designed by the Social Perceptional Lab of the TU Delft [4], nevertheless, this sensor was inspired by a similar sensor built by the MIT [19] and the Bluetooth settings of a normal smartphone. Therefore it is possible to work with differently recorded RSSI values if the relationship between the signal strength and the distance is known.

References

[1] V. Sekara and S. Lehmann, "The strength of friendship ties in proximity sensor data," in *PLoS ONE 9(7): e100915*, 2014. DOI: oi.org/10.1371/journal.pone.0100915.

[2] L. Cabrera-Quiros, A. Demetriou, E. Gedik, L. van der Meij, and H. Hung, "The matchnmingle dataset: A novel multi-sensor resource for the analysis of social interactions and group dynamics in-the-wild during free-standing conversations and speed dates," *IEEE Transactions on Affective Computing*, vol. 12, no. 1, pp. 113–130, 2021. DOI: 10.1109/TAFFC.2018.2848914.

[3] E. Gedik, "Capturing human behaviour through wearables by computational analysis of social dynamics [dissertation (tu delft)].," 2018.

[4] *Github - tudelft-spc-lab/spcl midge hardware: Hardware design files for the spc lab midge wearable devices for in-the-wild social data collection.* https://github.com/TUDelft-SPCLab/spcl_midge_hardware, 2019.

[5] H. Hung and B. Krose, "Detecting f-formations as dominant sets," Nov. 2011, pp. 231–238. DOI: 10.1145/2070481.2070525.

[6] Y. Hirabe, M. Fujimoto, Y. Arakawa, H. Suwa, and K. Yasumoto, "Effect on group detection based on human proximity for human relationship extraction in daily life," in *2017 46th International Conference on Parallel Processing Workshops (ICPPW)*, 2017, pp. 1–7. DOI: 10.1109/ICPPW.2017.14.

[7] R. Koshiba, Y. Hirabe, M. Fujimoto, H. Suwa, Y. Arakawa, and K. Yasumoto, "Group detection based on user-to-user distance in everyday life for office lunch group recommendation," in *2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA)*, 2017, pp. 309–314. DOI: 10.1109/WAINA.2017.43.

[8] S. Liu, Y. Jiang, and A. Striegel, "Face-to-face proximity estimation using bluetooth on smartphones," *IEEE Transactions on Mobile Computing*, vol. 13, no. 4, pp. 811–823, 2014. DOI: 10.1109/TMC.2013.44.

[9] Q. Lin and J. Son, "Analysis of bluetooth rssi for proximity detection of ship passengers," *Applied Sciences*, vol. 12, p. 517, Jan. 2022. DOI: 10.3390/app12010517.

[10] D. Leith and S. Farrell, “Coronavirus contact tracing: Evaluating the potential of using bluetooth received signal strength for proximity detection,” *ACM SIGCOMM Computer Communication Review*, vol. 50, pp. 66–74, Oct. 2020. DOI: 10 . 1145 / 3431832 . 3431840.

[11] M. Pavan and M. Pelillo, “Dominant sets and pairwise clustering,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 1, pp. 167–172, 2007. DOI: 10.1109/TPAMI.2007.250608.

[12] M. Swofford, J. C. Peruzzi, M. Vázquez, R. Martın-Martın, and S. Savarese, “DANTE: deep affinity network for clustering conversational interactants,” *CoRR*, vol. abs/1907.12910, 2019.

[13] Q. Dong and W. Dargie, “Evaluation of the reliability of rssi for indoor localization,” in *2012 International Conference on Wireless Communications in Underground and Confined Areas*, 2012, pp. 1–6. DOI: 10.1109/ICWCUCA.2012.6402492.

[14] F. Mirzaei and R. Manduchi, “Interpersonal proximity detection using rssi–based techniques,” *Indoor Positioning and Indoor Navigation (IPIN 21) - Work-In-Progress papers*,

[15] Z. Jianyong, L. Haiyong, C. Zili, and L. Zhaohui, “Rssi based bluetooth low energy indoor positioning,” in *2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2014, pp. 526–533. DOI: 10.1109/IPIN.2014.7275525.

[16] J. Huang, S. Chai, N. Yang, and L. Liu, “A novel distance estimation algorithm for bluetooth devices using rssi,” pp. 379–381, 2017. DOI: 10.2991/caai-17.2017.86.

[17] *Scipy - scipy.interpolate.interp1d*, https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.interp1d.html.

[18] B. Dikker, “Finding f-formations using dominant sets in the bluetooth proximity data of the conflab data set,” *TU Delft*, 2022.

[19] O. Lederman, D. Galacci, A. MacMullen, D. Fehder, F. Murray, and A. Pentland. “Open badges: A low-cost toolkit for measuring team communication and dynamics.” (2016), [Online]. Available: https://dam-prod.media.mit.edu/x/2019/01/10/IN_105.pdf.

A Result table

Table 6: The following table shows all results of experiment 1 and 2 conducted together. All possible window sizes were combined with all possible kernel sizes for each filter. This table should serve to get a deeper understanding how well each filter works with different parameter values. These values can be used in the future to tweak filters more.

Precision	Recall	F1-score	T	Window size	Kernel size	Filter
0.718	0.493	0.5854	1	20 sec	0	No filter
0.92	0.6499	0.763	2/3			
0.680	0.459	0.54	1	80 secs		
0.911	0.637	0.749	2/3			
0.6428	0.444	0.525	1	120 secs		
0.88	0.6250	0.734	2/3			
0.609	0.419	0.497	1	180 secs		
0.867	0.6022	0.71	2/3			
0.5646	0.3940	0.464	1	240 secs		
0.845	0.60	0.70	2/3			
0.719	0.4999	0.5899	1	20 sec	5	Median
0.924	0.6460	0.760	2/3			
0.67	0.46505	0.55	1	80 secs		
0.91	0.635	0.748	2/3			
0.64	0.447	0.5269	1	120 secs		
0.88	0.628	0.736	2/3			
0.611	0.420	0.497	1	180 secs		
0.867	0.609	0.7156	2/3			
0.56	0.395	0.464	1	240 secs		
0.84	0.599	0.701	2/3			
0.7269	0.496	0.59	1	20 sec	10	
0.927	0.66	0.7717	2/3			
0.667	0.507	0.576	1	80 secs		
0.91	0.69	0.788	2/3			

0.648	0.481	0.552	1	120 secs		
0.898	0.676	0.771	2/3			
0.627	0.43	0.512	1	180 secs		
0.883	0.617	0.727	2/3			
0.551	0.4253	0.48	1	240 secs		
0.849	0.63	0.723	2/3			
0.850	0.513	0.602	1	20 sec	20	
0.851	0.655	0.7675	2/3			
0.852	0.509	0.582	1	80 secs		
0.91	0.6929	0.786	2/3			
0.6469	0.48	0.551	1	120 secs		
0.895	0.6735	0.768	2/3			
0.629	0.437	0.5162	1	180 secs		
0.879	0.633	0.7367	2/3			
0.564	0.4181	0.480	1	240 secs		
0.861	0.625	0.7249	2/3			
0.7336	0.50	0.598	1	20 sec	30	
0.9260	0.6522	0.765	2/3			
0.504	0.6692	0.575	1	80 secs		
0.910	0.681	0.779	2/3			
0.646	0.4766	0.5487	1	120 secs		
0.8934	0.656	0.756998	2/3			
0.6245	0.431	0.51	1	180 secs		
0.88	0.634	0.7372	2/3			
0.554	0.424	0.480	1	240 secs		
0.8539	0.634	0.727	2/3			
0.73099	0.50	0.6007	1	20 sec	5	Gaussian
0.922	0.6544	0.765	2/3			
0.6674	0.503	0.574	1	80 secs		

0.90	0.685	0.78	2/3				
0.4750	0.6436	0.5466	1	120 secs			
0.895	0.661	0.7608	2/3				
0.625	0.43	0.5103	1	180 secs			
0.8810	0.622	0.729	2/3				
0.556	0.424	0.481	1	240 secs			
0.852	0.633	0.7267	2/3				
-	-	-	1	20 sec	10		
-	-	-	2/3				
0.675	0.511	0.58187	1	80 secs			
0.90896	0.69	0.78513	2/3				
0.6404	0.4769	0.546	1	120 secs			
0.892	0.664	0.761	2/3				
0.62	0.429	0.508	1	180 secs			
0.878	0.619	0.726	2/3				
0.56	0.423	0.482	1	240 secs			
0.852	0.63	0.7275	2/3				
0.7361	0.555	0.6333	1	20 sec	20		
0.921	0.6992	0.7953	2/3				
0.67	0.517	0.584	1	80 secs			
0.908	0.700	0.791	2/3				
0.653	0.477	0.551	1	120 secs			
0.894	0.65802	0.7584	2/3				
0.626	0.4321	0.5115	1	180 secs			
0.878	0.6167	0.7248	2/3				
0.5507	0.42	0.481	1	240 secs			
0.852	0.642	0.732	2/3				

0.72146	0.563	0.633	1	20 sec	30	
0.9269	0.717	0.808	2/3			
0.683	0.517	0.589	1	80 secs		
0.906	0.685	0.78	2/3			
0.653	0.477	0.55	1	120 secs		
0.893	0.657	0.757	2/3			
0.616	0.431	0.507	1	180 secs		
0.878	0.623	0.729	2/3			
0.558	0.435	0.489	1	240 secs		
0.85	0.644	0.7340	2/3			
					Cut-off frequency	
-	-	-	1	20 sec	0.02	Low pass filter
-	-	-	2/3			
0.71679	0.56611	0.632	1	80 secs		
0.9138	0.714	0.802	2/3			
0.7012	0.532	0.605	1	120 secs		
0.903	0.688	0.781	2/3			
0.65944	0.477	0.554	1	180 secs		
0.8913	0.641	0.746	2/3			
0.6011	0.4639	0.5236	1	240 secs		
0.874	0.644	0.742	2/3			
0.72580	0.575	0.6417	1	20 sec	0.07	
0.920	0.737	0.818	2/3			
0.70009	0.54625	0.6136	1	80 secs		
0.914	0.7102	0.799	2/3			
0.676	0.5024	0.5765	1	120 secs		
0.901	0.669	0.768	2/3			
0.6343	0.4458	0.5236	1	180 secs		

0.883	0.632	0.737	2/3			
0.579	0.43644	0.4978	1	240 secs		
0.863	0.636	0.732	2/3			
0.726	0.5717	0.64	1	20 sec	0.1	
0.9224	0.7277	0.8135	2/3			
0.694	0.5312	0.601	1	80 secs		
0.912	0.705	0.795	2/3			
0.6667	0.4998	0.571	1	120 secs		
0.901	0.673	0.77	2/3			
0.6316	0.4464	0.5231	1	180 secs		
0.884	0.633	0.7381	2/3			
0.5677	0.4377	0.494	1	240 secs		
0.865	0.639	0.735	2/3			
0.723	0.57	0.637	1	20 sec	0.3	
0.9287	0.9287	0.815	2/3			
0.6809	0.52750	0.594	1	80 secs		
0.909	0.7002	0.791	2/3			
0.662	0.4867	0.561	1	120 secs		
0.9031	0.6679	0.767	2/3			
0.6278	0.43756	0.515	1	180 secs		
0.8823	0.62859	0.7341	2/3			
0.5555	0.432	0.486	1	240 secs		
0.856	0.634	0.729	2/3			