Effects of Changing the Midge's Sampling Frequencies on Battery Life and Storage

J.V.P. de Kort
Supervisor(s): H. Hung, S. Tan, J. Vargas-Quiros
EEMCS, Delft University of Technology, The Netherlands

June 19, 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

The Socially Perceptive Computing Lab (SPCL) at Delft University of Technology has developed a device called the Midge. The aim of this device is to record data of social interactions at conferences. This paper aims to characterise how battery life is affected by different sensor settings on the Midge as well as how much data is generated in a given period of time. Based on the findings from the performed tests, formulae for both run time and data generation were devised.

1 Introduction

The Midge is a device largely based on the Open Badge [1] developed by the Human Dynamics group at the Massachusetts Institute of Technology. Like the Open Badge, the Midge also contains multiple modules including a Bluetooth Low Energy (BLE) transceiver and a microphone, however, in addition to this the Midge is also equipped with an inertial measurement unit (IMU). The Midge locally stores the telemetry collected from these sensors such that members of the SPCL can later use this data to analyse human behaviours in social settings.

Similarly to the Open Badge[1], analysis of group behaviour can be done through smartphones as seen in [2]. However, there are some key differences coming from the fact that the midge is specifically designed to analyse group behaviour with privacy in mind. First of all, it does not require participants of the study to continuously run an application on their private phones. A second difference between a phone and a Midge is the fact that it is designed to be worn around a person's neck with a lanyard, which should result in better audio recordings than a phone which is usually stashed away in a pocket or bag. Furthermore, the Midge is designed as a low-power device allowing it to record social group behaviour for long periods of time.

The Midge allows for some of its sensors to run with different sampling frequencies. Running with higher frequencies will generate more data and consume more power. The SPCL has not yet tested what effects this changing these parameters has. Therefore, this paper aims to characterize the impacts of different sampling and scanning frequencies of the Midge's sensors on battery life and data storage usage. And will answer the research question: "What are the maximum sampling frequencies the Midge's sensors can operate at whilst not exceeding data storage and battery limitations in a given time frame?" Such that the SPCL can decide what sampling frequencies they can use if they know the length of an event.

2 Methodology

The performed experiments aim to gain insights on how the battery life is affected as well as how much data is being written depending on the different sampling frequencies the Midge is set to record at. Although during testing both battery and data writing tests are done simultaneously, first the battery testing methodology is described in 2.1 whilst the storage-related testing is described in 2.2

2.1 Battery Testing Methodology

There exist multiple methods to analyse battery performance for different sensor sampling frequencies as described by [3]. Furthermore, one can disconnect the battery and replace it with a power supply to measure how much current the device draws at any given point. Even though this can give a quick view on how much power different sampling frequencies draw, the added complexity added by the need for desoldering the batteries and the need to still test how the batteries perform when the measured amount of power is drawn from them calls for a simpler and perhaps more realistic test like a discharge test.

In a discharge test, the battery is fully charged and then the device is run until it runs out of charge. Although running a Midge from fully charged to discharged is a time-consuming task it is relatively easy to set up and provides a realistic scenario which can be used to analyse the battery performance. Due to the realistic results it provides and the relative simplicity of performing a discharge test, it was decided to use discharge testing as the main type of test to analyse the Midge's battery performance.

External Influences on Testing

To accurately establish a baseline, first two types of discharge tests need to be performed. The first type focuses on assessing how external factors can influence a discharge test. To measure the influence the Midges are run on the default sampling settings. With these settings the Midges are then run in different scenarios, each scenario focusing on using a specific sensor and characterizing its effects on battery performance. The first scenario is testing Midges in which the Midges are almost constantly exposed to external sound and comparing the run times to Midges that have been in a mostly silent room, this test aims to verify that external sound inputs do not have an effect on battery life. Similarly, to see how the IMU's power usage is affected by motion, the second scenario is running tests with Midges in motion and comparing it to Midges laying motionless on a table. Finally, the third scenario consists of a single Midge running discharge tests compared to a Midge running discharge tests surrounded by 7 other active Midges. This final test aims to characterise the effect on battery life the Bluetooth module has when it scans doesn't find any nearby Midges compared to when it scans and does find a Midge. The results of these so-called scenario tests are described in the Results section. From these results, it can be concluded that none of the three external factors, sound, motion, and proximity have a significant influence on the battery life and therefore all further testing is performed with discharge tests without any special testing setup.

Changing IMU Frequencies

The IMU has a wide variety of settings. The first setting that can be changed is the data rate. This is a single setting that controls the sampling rate for the accelerometer (ACC), magnetometer (MAG), gyroscope (GYR) and rotational vector (ROT). An additional layer of complexity, as described in the IMU's manual [4] is the fact that the minimum and maximum sampling rates are 1Hz and 228Hz respectively for all 4 except for the minimum sampling rate of the ROT which is 56Hz and the maximum sampling rate for the MAG which is 76Hz. Setting the data rate setting outside the minimum or maximum possible frequencies will result in the sampling frequency for that specific sensor being capped to its possible minimum or maximum frequency. To measure the change in run time, Midges are run at their lowest (1 Hz) and highest (228 Hz) sampling frequency as well as with the IMU disabled and then compared to the baseline runs which were performed with a data rate of 50. The two other settings that can be changed for the IMU are the full scale range (FSR) of the gyroscope which can be set to 250, 500, 1000 or 2000 and the FSR of the accelerometer which can be set to 2, 4, 8 or 16. To see how the FSR affects the battery life, the FSRs are set to their maximum as well as their minimum settings in different tests and compared to each other.

Changing Microphone Settings

The Microphone on the Midge has 3 different settings, off, low frequency, and high frequency. To analyse the effects on battery life, Midges with different microphone settings but identical IMU and Bluetooth scanning settings are run and compared to each other. Furthermore, the microphone can be set to record in either mono or stereo mode. The SPCL indicated that mono sound recording is likely going to be used in the majority of the use cases therefore, with the exception of a few specific cases to evaluate stereo recording, the majority of tests will be performed with the Midges set to mono recording.

Changing Bluetooth Settings
The final sensor which requires testing is the Bluetooth Low Energy module (BLE). Unlike IMU and Microphone, this sensor does not have a limited amount of settings. It has two parameters, the scan interval and the scan window. The scan window defines how long the radio is active during a period. The scan interval defines the size of the window. With the limited testing time in mind, a small amount of values is chosen to get an understanding of how these two variables influence the battery life.

2.2 Storage Testing Methodology

There are two possible methods to deduce how much data is generated by a Midge. The first method is by looking at the specifications of a Midge and from those calculate the amount of data each sensor should generate in a specific amount of time. The second method is by running Midges at different settings and looking at the generated file sizes. After this, the file size and run time can be taken to calculate the amount of data generated per time period. Given the fact that multiple runs with different settings are performed for the battery testing, the run times and file sizes can be taken from these tests without the need for additional testing. On top of this, it will result in more real-world results. Therefore all file sizes from the battery tests will be recorded and used to evaluate how much data a Midge generates at different settings.

3 Test Setup

This section describes the test setup used for the different tests that were performed such that the experiments can be precisely repeated. In order to decide on how experiments would have to be set up, first experiments are necessary to find how external factors influence the Midges. Due to the lack of a temperature-controlled location, each test was performed in a normal room which means that the temperature may have slightly changed during and in between testing meaning the results may be influenced by temperature changes [5]. However, the slight variations were deemed negligible as well as unavoidable. Another variable which has to be taken into account is the fact that the run time of each Midge is affected by its battery capacity. Given the fact that the capacity of batteries can be impacted by factors like production imperfections, storage temperature, and the amount of charge/discharge cycles it has endured[6], it is crucial to keep track of which Midge gave which results and to only compare run times to those of itself.

3.1 Base Case

To get a baseline on the performance of each midge, all Midges were fully charged, and set to the same settings 1. Furthermore, each SD card was wiped of all old test data. After this, they were all turned on and set to record data and placed flat on a table. All Midges were placed in a row each separated by approximately 5 centimetres. The room was kept silent as much as possible during the period it took for all Midges to run out of battery.

3.2 Sound Test

This test was created to test whether or not external sounds have an impact on the size of the audio data recorded by the Midges, and to see if the run time is in any way affected. For this test, the Midges were set to their default settings 1 and placed in a circle around a laptop (HP Pavilion Gaming 15-cx0205ng) with a built-in speaker (L20345-001). This laptop was then set to continuously play a podcast [7] at 70% volume which translated to an average of 60DB when measured near the microphone of the Midges. This volume level was chosen because it is the average speaking volume[8]. The db measurement was performed with an application [9]running on an Iphone X. As soon as the fully charged Midges were set to record, the Podcast was started. The podcast was played continuously for 6 hours and 37 minutes in which it fully played episodes 109 through 114 in descending order. Unfortunately due to the room being used also being a bedroom, the playback had to be paused during the night. After the pause of exactly 6 hours and 20 minutes, the podcast was once again set to play during which it continuously played for 15 hours and 23 minutes from episodes 108 to 94. After this another pause was necessary and roughly 6 hours into the pause all midges ran out of battery. This means the midges were exposed to sound for exactly 22 hours which, depending on the midge is about two-thirds of its total run time.

3.3 Proximity Test

To test what effect the proximity detection of other Midges has on the battery life and size of the proximity file of the Midge, the baseline tests were used. As described in the baseline test setup, multiple Midges were run in close proximity to one another. The baseline tests alone do not reveal anything specific about how battery life or data generation change as all Midges ran in the same conditions, therefore additional tests were added in which a single Midge in isolation is run.

3.4 Motion Test

This test was performed to determine if the IMU detecting motion affects the run time of a Midge. To make the test as realistic as possible 4 Midges were placed in a bag which was attached to a lanyard. This lanyard was then worn by a person for the entire duration of the test, apart from 8 and 30 minutes during the night when the Midges were placed on a table whilst the person slept. The randomness of the person's motions throughout the roughly one and a half day period means the test is not precisely repeatable, however, to ensure that different types of motion do not result in different results, 4 more Midges were taped together and worn similarly to how an ankle bracelet is worn for the same amount of time as the Midges attached to the lanyard. Finally, this test was also separately repeated with 3 Midges attached to a lanyard. The different random movements of all 3 sets of midges ensure different motions were recorded by the different sets of Midges. The run times were then compared to those of the base case in which the Midges were motionless for the entire run time.

3.5 General Tests

Based on the results retrieved from the sound, proximity and motion tests, it is decided that all tests concerning Midge battery life and file sizes do not require a special setup in order to generate a more realistic scenario. Therefore each test was performed the same way as the baseline tests. These so-called discharge tests were performed for various different sensor settings.

3.6 Retrieving Test Data

The most accurate way to determine the total run time of a Midge would be to start a timer as soon as the Midge starts recording and to stop it as soon as the recording indicator LED turns off. This however would require constantly observing the LEDs of all the Midges which is not practical. Fortunately the telemetry the Midge records includes timestamps. Therefore in order to find the run time of a test, all that is required is to parse one of the binary files created by the Midge into a CSV file which then allows the first and last recorded timestamps to be found. Subtracting the timestamps of each other then results in the total run time. This method allows for easy retrieval of the run time, however, it is slightly less accurate than actively observing the LEDs because a Midge does not instantly write recorded telemetry to the SD card. Instead, it stores this telemetry in flash memory until either the flash memory is full or 5 minutes since the last flush to the SD card was made have passed and then flushes the telemetry to the SD card. This means that in some cases up to 5 minutes of time could be missing from the run time depending on when the battery runs out. a difference of up to 5 minutes is negligible compared to the total run time and the fluctuations in run time between identical tests in general.

4 Results

4.1 Baseline

The Baseline tests are tests that were performed to get a baseline for the Midges. To use the baseline to asses changes in run time, it is necessary to establish a separate baseline for each Midge since the battery capacity of each Midge can differ[6]. The results of these tests 2 show that when set to identical settings 1, the average run times of the used Midges are spread over 4 hours and 22 minutes. This fact highlights the need to only compare run time tests proportionally to other run time tests performed on the same Midge. Furthermore, the test results show that over the series of baseline tests, separately each Midge has a standard deviations ranged between 11 to 47 minutes when looking at run time per Midge. When determining a standard deviation of a set of results, ideally a lot more tests would be used, however time limitations restricted the ability to perform additional tests. What can however be done is to take the average of the standard deviations to get a better understanding. To find the average (Avg) of multiple standard deviations (STD) of unequal sample sizes one can use the following formula:

$$\text{Avg. STD} = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2+...+(n_k-1)s_k^2}{n_1+n_2+...+n_k-k}}$$

Where: n_k is the sample size of the k^{th} Midge.
s_k is the STD for the k^{th} Midge.
K is the total number of Midges.

Which, for the baseline tests results in an average sample deviation of 28 minutes. However, given the fact that some Midges last longer than others, it is more valuable to look at the normalised average STD which is 1,52%.

When performing the baseline tests, all file sizes were also recorded. These file sizes 3, are logically almost identical over equal periods of time. This means that unlike for run time testing it is not necessary to look at individual Midges. On default settings 1, the four files generated by the IMU for the ACC, MAG, ROT and GYR are identical in size and therefore displayed as one under the name IMU. Taking the average IMU file size divided by the run time gives an average of 0.08153MB per minute and a standard deviation 0.00046MB. Similarly the amount of Audio data generated by a Midge set to default settings is on average 0.14956MB with a standard deviation 0.00051MB. Unlike the IMU and Microphone, which have a very consistent data output as can be seen from the standard deviation, the proximity data recorded by the BLE module is a lot less consistent with an average of 0.00555MB and a standard deviation of 0.00127MB. This is due to a combination of 3 factors. BLE advertising makes use of a random interval to lower the probability of collisions [10]. This does not mean that there won't be any collisions and depending on the amount of nearby Midges and their BLE settings, the size of the proximity file may differ. Depending on the scan interval and window sizes as well as the random variable, two Midges may also not always find each other. The final factor has to do with the variation found in run times, namely if some Midges run out of battery earlier than in a test, there will be fewer Midges to detect for the Midges still running and therefore less data to write to the proximity file.

4.2 Sound Tests

To ensure that external sounds do indeed not influence the run time of a Midge, two sound tests were run. The results of the first run 4 show differences between 1.25 and 6.66% compared to the average baseline time set and all run times were longer than the baseline tests. After this run, an observation was made that the laptop gives off a considerable amount of heat, and given the close proximity of the heat exhaust vents to some of the Midges, the hot air may have affected battery life [5] At a later date the laptop was run for 30 minutes after which the temperature of the locations where the Midges had been, were measured. These locations were all between 3 and 7 degrees Celsius warmer than before the laptop was turned on. Furthermore the laptop may have slowly raised the average room temperature over the course of the entire test. In an effort to reduce this additional factor, for the second sound test special care was taken to try to shield the Midges from the laptop's heat. As can be seen from the results of the second test 5 the difference for each Midge to its average baseline test was no more than 0.83%. On top of this, due to a problem with Midge 38 which disabled the microphone, it was also recording during both audio tests without being affected by sound. In the second run where Midge 38 was shielded from the heat of the laptop, it lasted 55 minutes shorter, which, since it could not have been affected by the sound, could be another indicator that the discrepancy is heat related. Therefore, based on the second run only lasting 0.21% shorter than the baseline it is concluded that exposure to sound inputs does indeed not affect the run time. Similarly, comparing the audio file sizes of the audio tests 6 which show an average of 0.14956MB/minute and the baseline tests 3 which show an average of 0.14978MB/minute, it is clear that file size is not affected either.

4.3 Proximity Tests

The proximity tests compare Midges in isolation against the baseline tests where each Midge was in Bluetooth range of 6-7 other Midges. The run time tests of Midges in isolation 7 only consist of 4 measurements as due to the nature of this test no other Midges can be run within range meaning other Midges can not be tested at the same time. The results show that on average the run time only deviates by 0.20% from the baseline tests. From this, it can be concluded that having multiple Midges constantly detect each other over BLE does not significantly influence the run time. What is influenced however, is the amount of data generated, namely in every isolated run the Midges had a proximity file of size 0MB. This is logical as the Midge only writes to the proximity file when it detects another Midge. Given the fact that proximity does influence the storage tests one might suggest all further storage tests would require additional measures to have a constant proximity files size, however, given the fact that even with all available Midges running in close proximity on default settings, the biggest generated proximity file was 14.1MB which is negligible compared to the 4GB available storage on a Midge.

4.4 Motion Tests

In the first Motion test, Midges 16, 19 and 45 were worn around the ankle, whereas 18, 35, 39 and 49 were worn in a lanyard. The results of the test 8 provided an inexplicable range of results. Midges 16, 18 and 19 would suggest a negligible difference to the baseline whereas midges 35, 39, 45 and 49 on average differ 4.89%. Notably, this split between run time is not the same split as the one between where on the body the Midges were worn. A second run 11 in which all 3 Midges used were worn with a lanyard resulted in a 0.81% difference which also suggests that motion does not influence the results compared to the baseline. It is unclear why such a split exists in the first experiment. A possible explanation could have to do with the fact that the Midges were synced in increasing order of Midge number and the synchronisation of new settings may have accidentally been interrupted, however, no proof of this was visible. Based on the results of the second run it appears likely that motion does not influence the run time.

4.5 Sound Settings

The sound settings comprise of 3 options, namely off, low frequency and high frequency. On top of this, both the low and high frequency can be run in mono or stereo mode. Given the limited time for testing, most sound setting tests were centred around comparing the low frequency setting against the high-frequency setting with both set to mono sound recording as this will likely be the most used use case for the SPCL. From the first high frequency test, it became clear that no matter what the other settings are, if the Midge records high frequency audio it will always run out of storage before it runs out of battery. Assuming the Midges will be used with 4GB SD cards, this will always be the limiting factor. However, in case larger SD cards are used in the future, it is still useful to perform high frequency audio discharge tests that are not limited by storage. Therefore, all further high frequency tests were stopped around 11-18 hours into their run by turning off the Midge. After this, all data was removed from the SD card and placed on an external device after which the Midge was started again until it ran out of battery. Stitching the run times of the separate files together provides the following results 10. Comparing the run times to the baseline runs 2, it is visible that on average the high frequency run time is 4.34%. Turning off the microphone, as expected resulted in a longer run time. Compared to the baseline runs an average increase of 19.68% was measured15. When the Midge was set to low frequency and stereo recording 13 an average difference of 2.28% was observed. The change in battery consumption due to the stereo mode in low and high frequency, however, is not linear as compared to a mono high frequency run, the stereo high frequency run time was 7.01% shorter14.

The Microphone set to mono and low frequency sound generates 0.15MB per minute 3. As described in the previous section, the high frequency setting will always result in a Midge running out of storage space instead of battery. Even more so, the high frequency setting appears to be the only option capable of making the Midge run out of storage space before running out of battery. As shown in 16 on average the high frequency setting will generate 2.33MB of data per minute. Notably, compared to other data results, there is a significantly larger variation in the amount of data generated per minute. Switching to stereo recording for low frequency and high frequency generated 0.15MB 17 and 2.27MB 18 per minute respectively. Both these values fall within the 95 percent confidence interval of their respective mono settings. The microphone set to stereo recording does therefore appear to not generate more data than when set to mono.

4.6 Bluetooth Settings

Analysing the run time differences of the many different combinations of BLE window and interval settings that have been tested 19, there are 3 main observations. First, the upper and lower limit of run time that can be achieved by changing the variables lie around 19% increase and 12% decrease respectively. Furthermore, there is a clear trend between the ratio (RAT) of the window divided by the interval (INT) which is only broken by a few measurements, which leads to the third observation. The third observation is the fact that when two measurements have an identical ratio, the measurement with the smallest window and interval always results in a longer run time. Using only the data points with an interval of 1000, a trend line was extracted which follows the following formula:
% change in run time = $6.567 \times 10^{-1} \times RAT^2 - 9.634 \times 10^{-1} \times RAT + 1.949 \times 10^{-1}$. Similarly, a trend line: % change in run time = $-5 \times 10^{-5} \times INT + 5.8 \times 10^{-2}$ to characterise the effects of different interval sizes was extracted by only using measurements with a ratio equal to 0.25. Combining these two trend lines gives a formula that gives a rough estimation of the effects on run time 20.
Although the BLE module can have a severe impact on power consumption, it does not generate a lot of data. When disabled or when a Midge is out of Bluetooth range of other Midges, it does not write anything to storage. When there is only one other Midge within proximity an average of 0.40KB is recorded 21. Comparing this to the baseline tests 3 in which 5.55KB is recorded per minute, it can be seen that the size of the proximity file does not scale linearly to the number of nearby Midges. Given the fact that each Midge in the baseline tests had 7 midges within proximity for most of the time, it means that the baseline tests wrote twice as much data to the proximity file. Given the limited testing time, it was not possible to run all available Midges on the same BLE settings at the same time, therefore aside from the baseline test, no real meaningful data is available.

4.7 IMU Settings

Comparing the run time of Midges at the default IMU setting of 50Hz 2 to those with the IMU set to the highest possible sampling rate shows a difference of 5.96% 22. Looking at 22, 23 and 24 it appears that the FSR settings do not have an influence on run time as the different FSR settings result in similar run times. Furthermore, the results from 25 and 26 confirm that the two FSR settings do not cancel each other out. Based on these conclusions it may be sensible to take the average run time of the different IMU run times found in 22, 23 and 24 to get a better average for the run time with the IMU set to its highest setting. Changing the IMU to its lowest sampling frequency of 1Hz improved the baseline run times by 4.54% 27. When fully disabling the IMU, it becomes impossible to use the time stamps that are recorded in the IMU files. To get around this issue, the size of the audio recording was taken and divided by 0.15MB/minute which is the amount of data stored from the microphone per minute and gives the total run time of a Midge. From this an increase of 29.06% in run time was observed 28.

The IMU records the IMU data in 4 separate files, ACC, GYR, ROT and MAG. With the IMU set to its lowest sampling frequency of 1Hz, it writes 1.45KB of data per minute per file 29, with the exception of the MAG which is 81.30KB per minute as it cannot record below 56Hz. Trivially the amount of data written per minute scales linearly to the sampling frequency. This is can be confirmed by comparing 31 the average measured MB/minute for different frequencies 30 and the result of the 1Hz value multiplied by the corresponding frequency. The biggest deviation of 8.02KB is seen between the 1Hz value multiplied by 225 which translates to 14.7MB over a run time of 35 hours making it insignificant.

5 Problems and Limitations

While performing some tests, there were some problems and limitations. Here the effects of the problems and limitations are discussed in further detail as well as the actions taken to solve or alleviate them.

5.1 Ambient Temperature

A factor that can influence discharge tests is the ambient temperature. In an ideal case discharge tests should be performed with as little change in variables as possible, which would mean placing the Midges in a temperature-controlled room. Unfortunately this was not possible for the tests performed for this study. As discussed in the Results section, the ambient temperature may also have played a role in the sound test as the Midges were in close proximity to the running laptop which raised the ambient temperature and is likely the cause for the on average slightly higher run time.

5.2 Parsing Large Files

In order to find the start and end times of the discharge tests, it is necessary to parse the telemetry generated by the Midge into a human-readable format. For shorter runs the parser code provided by the SPCL does not provide a problem, however, parsing the large files generated by a full discharge test takes a long time and often resulted in the laptop crashing. Since only the first and last timestamp of one of the files generated is necessary it suffices to only parse one file, for instance, the accelerometer file. The parser code provides flags which are intended to disable the parsing of other files, unfortunately enabling the flags resulted in the parser crashing due to bugs in the code. To solve this issue, manual changes were made to the code such that it would no longer try to parse the other files.

5.3 Storage Limit

Each Midge comes equipped with 4GB SD cards. For all but one setting on the Midge, this is enough to perform a full discharge test without running out of storage capacity. The only problem when performing a discharge test appears when the Midge is set to record high-frequency audio. When this option is enabled, it will always run out of storage before running out of battery, which is an issue for performing a discharge test. To work around this issue without having access to a larger capacity SD card, the Midges set to high-frequency audio recording, are turned off roughly an hour before they would run out of storage. After this, the SD card is removed and its contents are copied to a computer and then wiped from the SD card. Finally, the SD card is placed back into the Midge and it is turned on and started again.

5.4 Bluetooth Proximity

The number of possible combinations of the two Bluetooth settings make it difficult to devise enough tests which can be used to get a good understanding of its effect on battery life and the number of data that will be generated. This becomes even more difficult with the added limitation of working with a small number of Midges, as in real use cases there may be many more Midges within range of a Midge which could mean more data would be generated than during the performed tests.

6 Responsible Research

This section reflects on the ethical implications with regards to the writing of this paper as well as the experiments performed. When researching it is important to be ethical and not to discard data that does not fit the hypothesis. For this reason no test results were discarded when they were not exactly in line with expectations. The exception to this are when one of the following 3 incidents occurred during testing:

- The physical audio switch being found to have accidentally switched during a test
- IMU data proving the sampling frequency did not match the intended sample rate
- Cases in which the settings were accidentally changed during a run when the intention was to connect to a different Midge.

In all 3 cases there is undeniable proof that the tests were not performed the intended way and were therefore removed. Another factor that could have been seen as an incident is the one described for the audio tests, where a notable change in temperature may have affected the results. For the sake of integrity, these results were not discarded but rather analysed and discussed in this paper.

Another issue is the issue of repeatability. To ensure other researchers can redo the experiments verify the results, the testing setup carefully describes how each test was performed. The setup section for example includes information of what speaker was used and exactly what audio was played. The only case which is not exactly repeatable is the motion test as it includes the randomness of human motions throughout the entire test, which due to the nature of the test was unavoidable.

Finally, since the Midges record data, privacy is also an issue. At any point during testing when another person entered the room, the person agreed to the fact that their voice could be recorded by a Midge. Similarly, in the case of the motion test, where the Midges were not in the room used for testing at all times, the people in proximity also agreed to the recording. Furthermore, since the goal of the experiments was to only characterise the amount of data and battery consumption, the data itself was deleted after the run time and file sizes were extracted.

7 Estimation Formulae

Using the data from the results, the following formulae were devised:

The effect the IMU has on run time:
$1 - (On(2 \times 10^{-6} \times IMU^2 - 1 \times 10^{-3} \times IMU + 4.62 \times 10^{-2}) + (1 - On) \times -2.977 \times 10^{-1})$
Where *On* is either 0 when the IMU is disabled and 1 when IMU is enabled and *IMU* is sampling frequency used by the IMU.

The effect the Microphone has on run time:
$7.67 \times 10^{-2} \times MIC^2 - 2.735 \times 10^{-1} \times MIC + 1.1968$
Where *MIC* is either 0 for off, 1 for low frequency and 2 for high frequency recording.

The effect of switching between mono and stereo has on run time:
$1 - MS \times (7.1 \times 10^{-3} \times MIC^2 + 2.09 \times 10^{-2} \times MIC)$
Where *MIC* is either 0 for off, 1 for low frequency and 2 for high frequency recording and *MS* is either 0 for mono recording and 1 for stereo recording.

The effect of changing the interval and window setting has on run time:
$Min((1 - (6.567 \times 10^{-1} \times (\frac{WIN}{INT})^2 - 9.634 \times 10^{-1} \times (\frac{WIN}{INT}) + 1.949 \times 10^{-1})) \times (1 + (5 \times 10^{-5} \times INT - 5.8 \times 10^{-2})), 1.19)$
Where *WIN* and *INT* are the scan window and interval respectively.

Due to the limited testing time, there have not been enough tests to validate the combinations of different possible settings when considering all 3 sensors at the same time, however multiplying the above 4 formulas with each other, as well as with the baseline time of an individual Midge set to the default settings 1 should give the estimated run time.

The amount of data in MB the IMU will generate per minute can be calculated with:
$2 \times 1.46 \times 10^{-3} \times IMU + Min(IMU, 76) \times 1.46 \times 10^{-3} \times Max(IMU, 56) \times 1.46 \times 10^{-3}$
Where *IMU* is the sampling frequency of the IMU. It should be noted that *IMU* is the real sampling frequency, of the IMU which is not always equal to the data rate setting which is used to set the IMU. For instance, when the datarate is set to 50, the real sampling frequency is 56Hz.

The amount of data in MB the Microphone will generate per minute can be calculated with:
$MIC \times ((2 - MIC) \times 1.5 \times 10^{-1} + (MIC - 1) \times 2.37)$
Where *MIC* is either 0 for off, 1 for low frequency and 2 for high frequency recording.

From the performed experiments, the amount of data generated by the BLE module can not be accurately estimated. Running 7 Midges in close proximity with default settings 1 generates 5.55×10^{-3} MB per minute. It is however not clear how this scales when more Midges are within Bluetooth range. For safety, it may be wise to reserve 5.55×10^{-2} MB per minute for the proximity file when using more Midges together until larger scale testing is done and this value can be adjusted.

Adding up the above two storage formulae and adding a margin for the proximity file will give the amount of MB per minute a Midge will generate.

8 Conclusions and Future Work

To characterise the run time, as well as the amount of data generated in a given time frame, the Estimation Formulae described in the previous can be used. With this, the main research question "What are the maximum sampling frequencies the Midge's sensors can operate at whilst not exceeding data storage and battery limitations in a given time frame?" can be answered. It should however be noted that due to limited scale and time of testing as well as uncontrollable outside influences during the performed experiments, the formulae are not perfect.

Bluetooth Settings

The Bluetooth testing results have shown that a considerable amount of run time can be gained or lost based on the Bluetooth radio settings. However, as discussed not all of these settings make sense to use for experiments, thus further research should be performed to find out what settings make sense for the SPCL to use after which more detailed discharge tests should be conducted. Similarly for data storage, in real experiments, there may be a lot more Midges within proximity which could make a notable difference in the amount of data generated. Larger scale testing should be performed with many more Midges to gain a better insight.

Sleep Modes

The SPCL indicated that they would like to use the Midges for multi-day events, however, based on the results from the experiments it is clear that with the current implementation the Midges will only last for two days at best. In order to achieve longer recording times without having to charge a Midge or clear its SD card, a sleep mode could be implemented. A mode like this could be created to stop recording during the night. Assuming that a person is for instance only at an event for 12 hours a day, then the midge could be in sleep mode for the other 12 hours and save a lot of battery and storage space by disabling the sensors.

Battery Degradation and Baselines For Each Midge

As discussed in multiple sections of the paper, each Midge will have a different run time depending on the state of its battery. Since the baseline performance of a Midge is used as a variable in the run time formula, each Midge used by SPCL should be tested to be able to calculate the maximum run time for each Midge. Assuming that social experiments done by SPCL require none of the Midges to run out before the experiment ends, the experiment time will be limited by the worst-performing Midge. To this end, SPCL could choose to only use the best performing Midges and gain a few hours, assuming that not all Midges in their possession are required for the experiment. Furthermore, since batteries degrade with each run [11] it may be useful to perform an analysis on how the battery life of a Midge degrades. Such an analysis could be used to update the baseline run times over time.

A Appendix

Setting Name	Setting Value
Scan_Window (ms)	250
Scan_Interval (ms)	1000
IMU_ACC_FSR (g)	4
IMU_GYR_FSR (dps)	1000
IMU_Datarate (Hz)	50
Microphone_Mode	1
Microphone_Frequency	Low

Figure 1: Default Midge Settings

IMU_ACC_FSR can be set to 2, 4, 8 or 16
IMU_GYR_FSR can be set to 250, 500, 1000 or 2000
Microphone mode can be set to 0 (stereo) or 1 (mono)
Microphone_Frequency can be set to Off, Low or High

Midge nr.	Run time 1	Run Time 2	Run Time 3	Run Time 4	Avg. Time	STD	STD/Avg. Time
16	34:18	33:37	33:38	33:38	33:47	00:20	0,99%
18	33:53	33:28	33:22		33:34	00:16	0,82%
19	32:38	32:13	32:12	32:28	32:22	00:12	0,64%
35	29:43	30:38	30:58	30:48	30:31	00:33	1,83%
38	29:03	29:48			29:25	00:31	1,80%
39	29:18	30:28	30:48		30:11	00:47	2,61%
45	30:32	31:23	31:33		31:09	00:32	1,75%
47	33:09	33:10	32:49		33:02	00:11	0,60%
49	29:43	30:03	30:48		30:11	00:33	1,84%
					Avg. STD:	**00:28**	**1,52%**

Figure 2: Baseline run time results

Run time is in HH:MM format
Avg. Time is the average run time over all baseline test per Midge STD/Avg. Time shows a normalised STD

Run Time	IMU Size	MB/Minute	Audio Size	MB/Minute	Prox. MB	MB/Minute
2018	164,6	0,08157	302,2	0,14975	13,0	0,00644
2002	163,6	0,08172	299,5	0,14960	08,0	0,00400
1932	158,3	0,08194	289,4	0,14979	07,8	0,00404
1858	150,4	0,08095	278,1	0,14968	08,6	0,00463
1788	147,3	0,08238	267,7	0,14972	07,4	0,00414
1848	150,0	0,08117	276,6	0,14968	09,4	0,00509
1893	153,7	0,08119	283,3	0,14966	11,0	0,00581
1830	148,2	0,08098	270,4	0,14776	10,0	0,00546
2058	167,9	0,08158	308,3	0,14981	08,0	0,00389
2033	166,4	0,08185	304,4	0,14973	09,5	0,00467
1958	160,4	0,08192	293,3	0,14980	08,6	0,00439
1783	144,5	0,08104	266,9	0,14969	09,1	0,00510
1758	142,8	0,08123	263,2	0,14972	08,4	0,00478
1832	149,0	0,08133	274,4	0,14978	08,4	0,00459
2017	164,4	0,08151	302,0	0,14973	12,5	0,00620
2008	164,2	0,08177	300,6	0,14970	12,6	0,00627
1933	158,3	0,08189	289,4	0,14972	14,1	0,00729
1838	148,8	0,08096	275,1	0,14967	12,7	0,00691
1743	143,6	0,08239	260,9	0,14968	14,1	0,00809
1828	148,3	0,08113	273,6	0,14967	12,6	0,00689
1883	152,9	0,08120	281,8	0,14965	13,7	0,00728
1783	146,6	0,08222	266,9	0,14969	11,7	0,00656
1989	163,0	0,08195	298,4	0,15003	07,3	0,00367
1990	160,9	0,08085	294,7	0,14809	13,6	0,00683
1969	160,4	0,08146	293,8	0,14921	11,3	0,00574
	Average:	**0,08153**		**0,14956**		**0,00555**
	STD:	**0,00046**		**0,00051**		**0,00127**

Figure 3: Baseline data size results

Run Time is in minutes
IMU Size is defined as the size of only the ACC file. Which is identical to the size of the GYR, ROT and MAG, therefore the total IMU size is 4 times bigger than in the table
Audio Size is the total size of all audio files combined
Prox. Size is the size of the proximity file

Midge nr.	Run Time	Diff to Base
16	34:22	1,69%
18	34:07	1,62%
19	32:47	1,25%
35	32:22	6,02%
39	32:12	6,66%
45	32:47	5,22%
49	31:32	4,45%
	Total Avg:	**3,85%**
	STD:	**2,28%**

Figure 4: Sound test 1 results

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr.	Run Time	Diff to Base
16	33:32	-0,78%
18	33:22	-0,61%
19	32:17	-0,30%
35	30:47	0,83%
	Total Avg:	**-0,21%**
	STD:	**0,73%**

Figure 5: Sound test 2 results

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Run Time	Audio Size	MB/Minute
2062	308,8	0,14976
2047	306,7	0,14983
1967	294,6	0,14977
1942	290,8	0,14974
1932	289,3	0,14974
1967	294,8	0,14987
1892	283,3	0,14974
2012	301,4	0,14980
2002	299,7	0,14970
1937	290,2	0,14982
1847	276,6	0,14976
	Average:	0,14978
	STD:	0,00005

Figure 6: Sound test 1 and 2 data size results

Run Time is in minutes
Audio Size is the total size of all audio files combined

Midge nr.	Run Time	Diff to Base
16	33:08	-1,96%
18	33:38	0,18%
18	33:33	-0,07%
19	32:43	1,04%
	Total Avg:	**-0,20%**
	STD:	**1,27%**

Figure 7: Proximity test results of Midges in isolation

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr.	Run Time	Diff to Base
16	34:02	0,70%
18	33:47	0,63%
19	32:38	0,78%
35	32:30	6,46%
39	31:32	4,45%
45	32:38	4,74%
49	31:22	3,90%
	Total Avg:	**3,10%**
	STD:	**2,37%**

Figure 8: Motion test 1 results

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr.	Run Time	Diff to Base
16	33:55	0,36%
35	30:43	0,61%
49	30:38	1,47%
	Total Avg:	**0,81%**
	STD:	**0,58%**

Figure 9: Motion test 2 results

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr	Run Time	Diff to base
18	31:41	-5,63%
18	32:05	-4,43%
45	29:41	-4,73%
45	30:21	-2,59%
	Total Avg:	-4,34%
	STD:	1,28%

Figure 10: High Frequency test results

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr.	Run Time	Diff to Base
16	33:55	0,36%
35	30:43	0,61%
49	30:38	1,47%
	Total Avg:	**0,81%**
	STD:	**0,58%**

Figure 11: Motion test 2 results

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr	Run Time	Diff to base
38	36:22:00	23,59%
38	34:42:00	17,93%
38	35:27:00	20,48%
35	35:38:00	16,72%
	Total Avg:	**19,68%**
	STD:	**3,04%**

Figure 12: Microphone off test results

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr	Run Time	Diff to Base
47	32:38	-1,24%
47	31:57	-3,31%
	Total Avg:	-2,28%
	STD:	1,46%

Figure 13: Stereo low frequency recording results

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr	Run Time	Diff to HF Base
47	29:23:00	-7,01%

Figure 14: Stereo high frequency recording results

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr	Run Time	Diff to base
38	36:22:00	23,59%
38	34:42:00	17,93%
38	35:27:00	20,48%
35	35:38:00	16,72%
	Total Avg:	**19,68%**
	STD:	**3,04%**

Figure 15: Microphone off test

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Minutes	Audio Size	MB/Minute
1901	4419,7	2,32493
1781	4139,7	2,32437
1925	4487,9	2,33138
1869	4245,4	2,27148
1444	3469,3	2,40256
1447	3464,4	2,39419
1145	2633,5	2,30000
1150	2654,5	2,30826
	Average:	**2,33215**
	STD:	**0,04505**

Figure 16: High frequency data size results

Run Time is in minutes
Audio Size is the total size of all audio files combined

Minutes	Audio Size	MB/Minute
1958	294,1	0,15020
1917	287,5	0,14997
	Average:	**0,15009**
	STD:	**0,00016**

Figure 17: Stereo low frequency data size results

Run Time is in minutes
Audio Size is the total size of all audio files combined

Minutes	Audio Size	MB/Minute
1823	4141,9	2,27202414

Figure 18: Stereo high frequency data size results

Run Time is in minutes
Audio Size is the total size of all audio files combined

Midge nr.	Window	Interval	Window/Interval	Run time	Diff to Base
19	0	0	0,000	38:18	18,29%
19	5	1000	0,005	38:30	18,90%
49	5	1000	0,005	35:58	19,14%
18	100	1000	0,100	37:08	10,61%
18	100	1000	0,100	37:09	10,66%
16	100	500	0,200	34:58	3,46%
39	25	100	0,250	31:59	5,94%
Baselines	250	1000	0,250	00:00	0,00%
38	500	2000	0,2500	34:28	-2,93%
19	300	1000	0,300	30:48	-4,88%
39	100	300	0,333	29:34	-2,06%
45	100	300	0,333	30:23	-2,48%
45	50	100	0,500	30:03	-3,55%
45	250	500	0,500	27:33	-11,57%
45	250	500	0,500	27:32	-11,63%
19	500	1000	0,500	28:38	-11,57%
49	250	250	1,000	26:58	-10,67%
49	250	250	1,000	27:08	-10,12%
49	1000	1000	1,000	26:58	-10,67%
49	1000	1000	1,000	26:38	-11,78%

Figure 19: BLE test results

Run Time is in HH:MM format
Baselines are the combined results of the baseline tests
Diff to Base is the (per Midge) % change in run time compared the baseline test

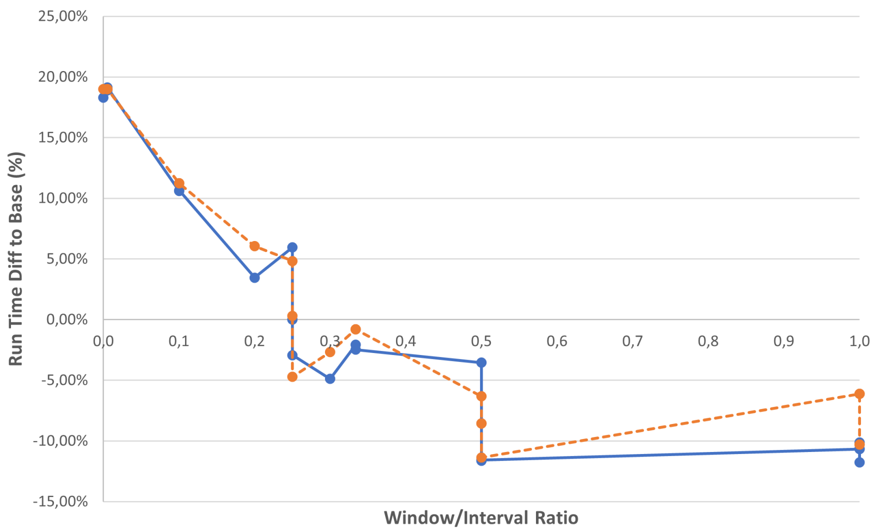

Figure 20: Real BLE setting results compared to estimated run times

Orange (Dashed) line indicates the predicted change in run time
Blue line indicates the actual run time

Minutes	Prox Size	MB/Minute
2018	0,782	0,00039
1958	0,794	0,00041
	Average:	**0,00040**
	STD:	**0,00001**

Figure 21: 2 Midges within proximity data size results

Run Time is in minutes
Prox Size is the size of the proximity file

Midge nr.	Run Time	Diff to base
19	30:30	-5,80%
19	30:24	-6,11%
	Total Avg:	-5,96%
	STD:	0,22%

Figure 22: Maximum IMU sample rate with default FSR settings results

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr.	Run Time	Diff to base
16	31:35	-6,55%
16	31:26	-6,99%
	Total Avg:	**-6,77%**
	STD:	**0,31%**

Figure 23: Maximum IMU sample rate with the highest FSR settings results

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr.	Run Time	Diff to Base
16	31:20	-7,29%
16	31:21	-7,24%
	Total Avg:	-7,26%
	STD:	0,03%

Figure 24: Maximum IMU sample rate with the lowest FSR settings results

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr.	Run Time	Diff to Base
35	30:48	0,89%

Figure 25: Normal IMU sample rate with ACC_FSR and GYR_FSR set to the lowest, and highest setting respectively

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr.	Run Time	Diff to Base
16	33:58	0,51%

Figure 26: Normal IMU sample rate with ACC_FSR and GYR_FSR set to the highest, and lowest setting respectively

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr.	Run Time	Diff to base
35	31:49	4,22%
35	31:59	4,76%
39	31:24	4,01%
39	31:45	5,17%
	Total Avg:	4,54%
	STD:	0,53%

Figure 27: Minimum IMU sampling frequency test results

Run Time is in HH:MM format
Diff to Base is the (per Midge) % change in run time compared the baseline test

Midge nr.	Audio Size	Run time	Diff to Base
35	355,5	39:36	29,77%
35	345	38:26	25,93%
39	356,2	39:41	31,49%
		Total Avg:	29,06%
		STD:	2,84%

Figure 28: IMU disabled run time test results

Run Time is in HH:MM format
Audo Diff to Base is the (per Midge) % change in run time compared the baseline test
Audio Size is the total size of all audio files combined
Run time was calculated using the Audio Size and an estimated 0.15MB/minute audio generation rate.

Run Time	IMU Size	MB/Minute	Rotation Size	MB/Minute
1909	2,8	0,001467	155,1	0,081247
1919	2,8	0,001459	155,9	0,081240
1884	2,7	0,001433	153,4	0,081423
1914	2,8	0,001463	155,5	0,081243
1905	2,8	0,001470	155,0	0,081365
	Average:	**0,001458**		**0,081304**
	STD:	**0,000015**		**0,000085**

Figure 29: Minimum IMU data size test results

Run Time is in minutes
IMU Size is defined as the size of only the ACC file. Which is identical to the size of the GYR and MAG, therefore the total IMU size is 3 times bigger than in the IMU value in the table
Rotation Size is the size of the ROT file

Run Time	IMU Size	MB/Minute	Mag. Size	MB/Minute
1895	614,7	0,32438	204,9	0,10813
1830	595,9	0,32563	198,6	0,10852
1886	606,1	0,32137	202,0	0,10710
1824	591,2	0,32412	197,1	0,10806
1880	601,9	0,32016	200,6	0,10670
1145	362,7	0,31677	120,9	0,10559
1150	362,6	0,31530	120,9	0,10513
	Average:	**0,32110**		**0,10703**
	STD:	**0,00395**		**0,00131**

Figure 30: Maximum IMU data size test results

Run Time is in minutes
IMU Size is defined as the size of only the ACC file. Which is identical to the size of the ROT and MAG, therefore the total IMU size is 3 times bigger than in the IMU value in the table
Mag. Size is the size of the MAG file

IMU Hz	Real	Estimated
1	0,00146	0,00146
56	0,08153	0,08167
76	0,10703	0,11083
228	0,32110	0,33250

Figure 31: Results of using the real 1Hz IMU file size to calculate files sizes of other IMU frequency file sizes

IMU Hz is the sampling frequency of the IMU
Real is the measured amount of MB per minute generated by the IMU
Estimated is the calculated amount of MB per minute generated based on the 1 Hz IMU value

References

[1] 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.

[2] S. Dash, S. K. Shakyawar, M. Sharma, and S. Kaushik, “Human behaviour analysis through smartphones,” *Proceedings (MDPI)*, vol. 2, October 2018.

[3] R. Rao, S. Vrudhula, and D. Rakhmatov, “Battery modeling for energy-aware system design,” *Computer*, vol. 36, pp. 77 – 87, 01 2004.

[4] T. Invensense, *Software User Guide For DK-20948 Dev Kit*, 2018.

[5] S. Lv, X. Wang, W. Lu, Z. Jiaqiao, and H. Ni, “The influence of temperature on the capacity of lithium ion batteries with different anodes,” *Energies*, vol. 15, p. 60, 12 2021.

[6] L. Shenzhen Pknergy Energy Co., *Li-Polymer Battery Technology Specification (LIPO503035 500mAh 3.7V)*.

[7] M. Brownlee, “Waveform.” `https://podcasts.apple.com/us/podcast/waveform-the-mkbhd-podcast/id1474429475`, 2022.

[8] U. S. Government, “What noises cause hearing loss?.” `https://www.cdc.gov/nceh/hearing_loss/what_noises_cause_hearing_loss.html`, 2019.

[9] Vlad Polyanskiy, “Decibel: db sound level meter.”

[10] Argenox_Technologies_LLC., “Ble advertising primer.” `https://www.argenox.com/library/bluetooth-low-energy/ble-advertising-primer/`.

[11] V. Ovejas and A. Cuadras, “Effects of cycling on lithium-ion battery hysteresis and overvoltage.,” October 2019.