



ActiGraph White Paper:

Scaling and Correcting Accelerometer Data Affected by 16CPAN03

Author List

NATHAN MILLER¹

nathan.miller@actigraphcorp.com

850-332-7900

49 E Chase St., Pensacola, FL 32502

JOE NGUYEN¹

Joe.nguyen@actigraphcorp.com

850-332-7900

49 E Chase St., Pensacola, FL 32502

W. JEREMY WYATT¹

Jeremy.wyatt@actigraphcorp.com

850-332-7900

49 E Chase St., Pensacola, FL 32502

¹ActiGraph, LLC

Abstract

In March of 2016, ActiGraph identified an issue that caused a data anomaly (raw data attenuation) on wGT3X-BT activity monitor units shipped between November 7, 2013 and March 24, 2014 which had collected data with firmware versions 1.4.0, 1.4.1, 1.5.0, 1.6.0 or 1.7.0. In order to minimize the impact on academic publications which rely on data from these units, ActiGraph issued 16CPAN03, a Corrective and Preventative Action (CAPA) case. The CAPA process included 1) notifying affected customers, 2) releasing a firmware update to permanently fix the issue and 3) providing file repair options to affected customers. ActiGraph developed three file repair alternatives: File Scanning, Calibration File Application, and Factory Calibration. In all three fix cases, the incorrect calibration applied to the data file is undone, new calibration constants are discovered, and the new constants are applied to the data file, thus fixing the file. The first two methods (File Scanning and Calibration File Application) require that the affected data file(s) meet strict criteria. With these methods, a Genetic Algorithm (GA) is employed to estimate the correct calibration constants. The third method (Factory Calibration) repeats the calibration procedure done at factory and applies the measured differences to correct the file. All three methods have shown effectiveness in fixing files.

Key Words: ActiGraph, Firmware, File Repair, CAPA

Introduction

Accurate observation and measurement of physical phenomenon is one of the keys to good science. In the case of objective activity monitoring, correctly measuring acceleration is crucial. Micro-Electrical Mechanical Systems (MEMS) based accelerometers have been widely employed due to their accuracy, durability, linearity, and price points. Although this technology has greatly improved in recent years, calibration is still required to achieve optimal accuracy.

Calibration

During the manufacturing process, a test jig is used to expose the unit-under-test to exactly one unit of earth's gravity (1G) in six different planes. The accelerometer digital outputs are measured from each of the three axes in each plane, and from those measurements, sensitivity (milliGs/Least Significant Bit) and zero G offset (in Gs) are calculated and recorded as static calibration coefficients. Note that the sensitivity calibration error should never exceed $\pm 10\%$ (from 1G); this behavior is expected due to the variation of inter-device sensitivity (4). These calibration coefficients are stored in nonvolatile memory and then, during normal device operation, are applied to every measurement sample prior to data storage. This calibration and normalization process helps ensure inter-device consistency and measurement accuracy.

16CPAN03 Issue Description

wGT3X-BT units manufactured prior to March 24, 2014 contained firmware that programmed the on-board MEMS accelerometer (ADXL352) with a specific set of antialiasing filter parameters prior to calibration. On March 24, 2014, a firmware update was released (publicly and to ActiGraph's contract manufacturer) for the wGT3X-BT with serial numbers beginning with MOS. This update changed those antialiasing settings in newly manufactured units in an

effort to improve the reduction of antialiasing of the MEMS accelerometer included in the wGT3X-BT. Because these changes slightly affected the calibration coefficients, units manufactured prior to March 24, 2014 were required to use the prior configuration indefinitely. The firmware released on March 24, 2014 (version 1.1.0) accounted for this change by essentially determining at run time which configuration needed to be selected based on the calibration date and then applying that configuration.

On March 11, 2015, a wGT3X-BT firmware update was released (version 1.4.0) that inadvertently removed the aforementioned run-time check. For units manufactured prior to March 24, 2014, this meant that the internal calibration constants did not match the accelerometer configuration. The result, and the subject of this paper, was a systematic, albeit slight, attenuation in raw accelerometry samples for all affected units.

On March 9, 2016, this issue was discovered. On May 16, 2016, a firmware patch was released (version 1.8.0) which reintroduced the configuration switch that was removed in version 1.4.0, thereby permanently fixing the issue for data collected henceforth (1). A timeline of events is summarized in *Figure 1*.

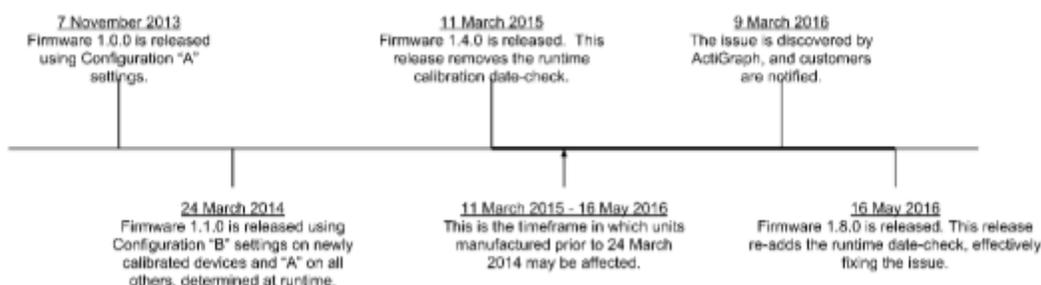


Figure 1- Timeline showing the sequence of events.

Effects on Raw Data

For units impacted by the 16CPAN03 issue, the attenuation of the raw data (raw accelerometry samples from the ADXL362) is apparent when looking at a device at rest, although the issue affects the entire range of possible data. The affected device resting on a desk with a major axis facing gravity does not report +1 or -1 gravity as expected. *Figure 2* shows data from a collection of units. Unaffected units average about 25 milligravity (mG) difference from the 1G expected reading, whereas units affected by the 16CPAN03 issue exhibit a difference of around 175 mG. Higher levels of acceleration display similar offset and gain issues.

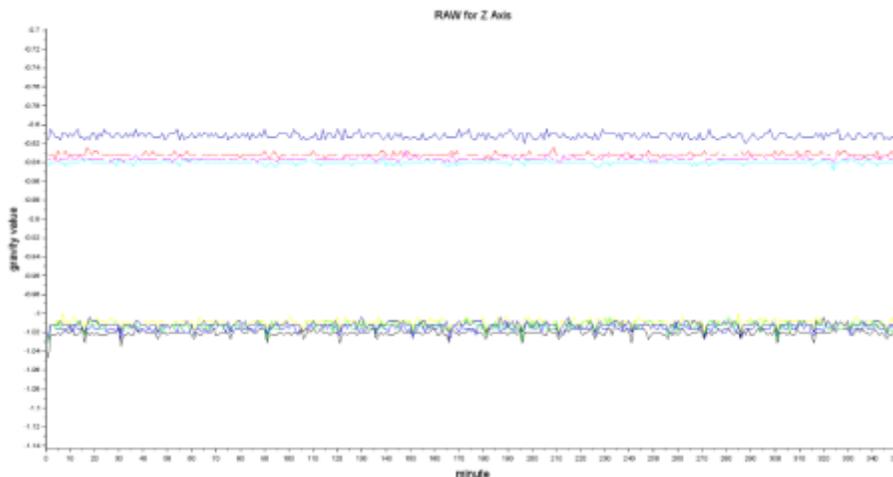


Figure 2 - Z axis measured accelerations. Good unit readings cluster near -1 earth gravity. Bad units cluster between -0.8 and -0.85 earth gravity

Effects on Post-Processed Data

ActiGraph's ActiLife software allows users to process accelerometer data into a myriad of various outcomes (energy expenditure, cut points, sleep efficiency) based on referenced academic publications. Many users reference these outcomes in their own published literature. It is important, therefore, to note the 16CPAN03 issue impacts these post-processed outcomes.

Methods

The Approach

During data collection, ActiGraph devices apply the calibration coefficients (stored in nonvolatile memory during the manufacturing process) to every data sample coming from the ADXL362 accelerometer to correct any offset and gain error inherent in the accelerometer and then store the sample in nonvolatile memory. For 16CPAN03 affected units, these calibration coefficients are incorrect. A process has been developed to discover correct calibration coefficients and correct the data. The following steps are performed when repairing affected files:

1. Reverse existing calibration by using the existing coefficients to undo the scaling process. Find reference data by scanning the file for periods of inactivity or by loading inactivity data from a specially created calibration file containing these periods of inactivity.
2. Compute new calibration coefficients by using a Genetic Algorithm (GA) to search for new coefficients using the reference data from step 3, or by performing the factory calibration procedure.
3. Verify that the new coefficients by examining the results of the GA.
4. Scale the data using the new coefficients to produce a corrected data file.

Process Variations

Three methods are proposed herein to repair data files produced by units which were impacted by the 16CPAN03 issue. Each method essentially aims to achieve the same goals: identifying and correcting for the error in the calibration coefficient constants.

Step 1: Reverse Existing Calibration

Accelerometers produce a linear output. Because the errant files contain raw data that has been scaled improperly, the first step in repairing those files is to reverse the scaling (i.e., the application of the existing calibration scale and offset factors to the raw data). The existing (incorrect) calibration coefficients are stored in the .gt3x data file, so they can easily be retrieved. It is then possible to remove calibration offset and gain from the data file by reversing the linear scale and offset to derive the original accelerometer data samples, thereby normalizing the file.

Step 2: Find Recalibration Points

Assuming the unit (from which the incorrect calibration constants can be directly measured) is physically unavailable, a repair method must be developed that can accurately discern the correct calibration from the errant raw data files. This is possible with autocalibration type algorithms such as is described in *Autocalibration of MEMS Accelerometers* and *Automated Antenna Design with Evolutionary Algorithms* (2, 3). Following the methods described therein, earth's gravity can be derived as a reference during empirical measurement "based on the assumption that an inertial sensor, in static condition, is subjected only to the gravity force." (2) Affected raw data files can be scanned for these reference periods of inactivity; then, using a Genetic Algorithm (GA) (discussed below), new calibration coefficients can be produced.

The calibration equations, as implemented in ActiGraph firmware, yielding the actual X, Y, and Z output are:

$$XSensitivity = 2 \div (XPositiveOffset - XNegativeOffset)[1]$$

$$YSensitivity = 2 \div (YPositiveOffset - YNegativeOffset)[2]$$

$$ZSensitivity = 2 \div (ZPositiveOffset - ZNegativeOffset)[3]$$

$$X_{scaled} = XSensitivity \times (X_{measured} - XZeroOffset) [4]$$

$$Y_{scaled} = YSensitivity \times (Y_{measured} - YZeroOffset) [5]$$

$$Z_{scaled} = ZSensitivity \times (Z_{measured} - ZZeroOffset) [6]$$

Where *PositiveOffset* is the positive offset of the axis, *NegativeOffset* is the negative offset of the axis, and *ZeroOffset* is the zero offset of the axis. *Sensitivity* is the computed sensitivity coefficient of the axis.

$$VectorMagnitudeGravity = \sqrt{X_{axis}^2 + Y_{axis}^2 + Z_{axis}^2} [7]$$

In the case of static inactivity points, *VectorMagnitudeGravity* = 1. This system of equations with nine unknowns can be solved using various techniques.

To identify the proper calibration coefficients in the corrupted files, a technique that exposes points on the best fit line equation (Equations 4, 5, and 6) must be developed. Our experiments have shown that if all axes (-X, +X, -Y, +Y, -Z, +Z) in the corrupted (and uncalibrated) files are exposed to more than 0.3G, computing calibration coefficients is possible using a Genetic Algorithm.

We have developed two methods for finding suitable inactivity data on which this method can be performed:

1. The first method is simply to scan the affected data file for periods of inactivity. If enough periods of inactivity are found in the affected file, we can run the GA on that data.
2. If there are not enough suitable periods of inactivity in the affected data file, but the unit that produced the data file is available, then the unit can be used to generate a

“calibration file,” or simply a data file collected while the device was placed on each side and allowed to rest for enough time that the inactivity data is produced. The process of generating the calibration file is thoroughly described on ActiGraph’s knowledge base (1).

The first step in finding the best fit line is identifying the periods of inactivity from which static acceleration can be measured. A legitimate period of inactivity from which calibration can later be derived is defined as any 4-second long block of data in which the following are true:

1. Linear regression on the block produces a slope less than 0.0005 g/sample.
2. Linear regression on the block produces a maximum error of all original points relative to regression line for the block is less than 0.015G;
3. For each data point: $\text{abs}(x) > 0$ and $\text{abs}(y) > 0$ and $\text{abs}(z) > 0$

The last requirement is due to the fact that the wGT3X-BT unit records all axes as having 0G acceleration when it is charging (and therefore not in use). Requiring a nonzero acceleration on at least one axis therefore ensures that data is selected from a period when the unit was recording data.

To ensure that the file is indeed in need of calibration, each inactivity data point should be checked to verify the absolute value of the vector magnitude sum is equal to $1 \text{ G} \pm 0.064 \text{ G}$. If any points deviate further than 0.064 G from 1G, the file is indeed corrupt and needs to be repaired. The 0.064G uncertainty is determined based on a combination of potential accelerometer noise and drift caused by temperature variation.

Inactivity locations that are found are grouped to into subsets for later use with the Genetic Algorithm. For each axis, there will be three value subsets that are grouped according to the gravity value at the location: +1.3g to +0.7g, +0.7g to -0.7g , and -0.7g to -1.3g. In total, there will be nine specific subsets for all three major axes. *Figure 3* illustrates a sample of inactivity

data showing points in each value-subset, with the valid inactivity period appearing in the -0.7G to 0.7G subset.

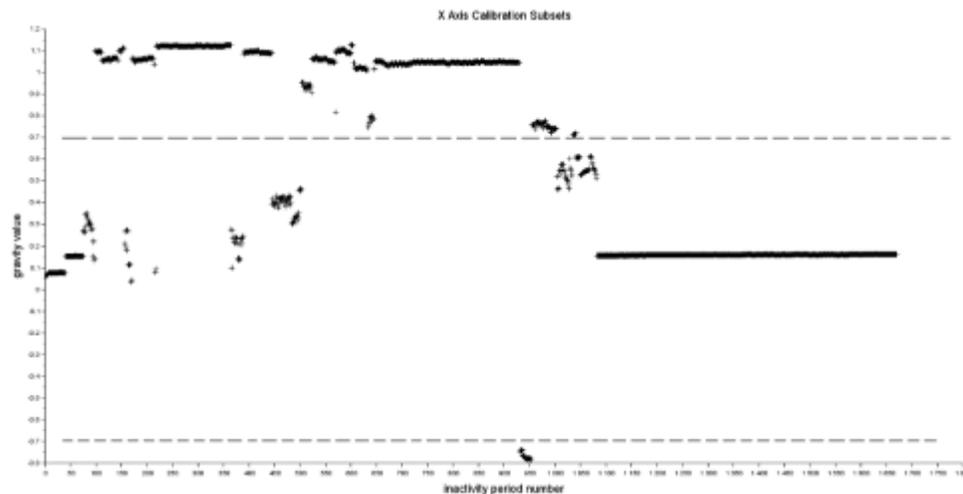


Figure 3 - A sample of inactivity data showing points in each value-subset, with the valid inactivity period appearing in the -0.7G to 0.7G subset.

Alternatively, if no suitable inactivity locations are found in the file and the device is available, a data file can be generated with suitable inactivity data that can be used in the next step (Method 2).

Step 3: Compute New Calibration Coefficients using Genetic Algorithm

There are two possible processes for generated new calibration coefficients for an affected unit.

If suitable periods of inactivity data are found in Step 2, then a Genetic Algorithm can be used to search for new calibration coefficients. Otherwise, if the unit is physically available, it can be returned to ActiGraph and have the factory calibration process repeated with the new filter settings.

As described in *A genetic algorithm tutorial* (6), the term genetic algorithm (GA) refers to a family of biological computational models developed by John Holland and discussed in his book *Adaptation in Natural and Artificial Systems*. These models have been applied to solving complex optimization problems in which searching for an optimal solution would be difficult using standard search methods such as brute force or Newton-Raphson. Newton-Raphson requires 2nd order derivatives, which may be unknown or hard to derive, for the system of equations to be solved. The size of the search space of a brute-force solution makes it infeasible, due to the large computational requirements. In the case of identifying the correct calibration coefficients, the GA method provides a clear means for performing what is essentially a nonlinear optimization for a function of nine variables. The correct calibration coefficients can be calculated by taking a subset of the re-calibration points data set and adjusting calibration coefficients so the fix points are as close as possible to 1 G vector magnitude (see equations in Step 2).

Now that the incorrect calibration has been removed and reference points have been identified (Steps 1 and 2), the GA approach can be used to estimate a new, correct set of calibration coefficients. The processes used by the GA to discover the new coefficients are described below.

Brief overview of the Genetic Algorithm

A GA searches for coefficients by modeling biological processes as described in the *A genetic algorithm tutorial* (6). For each iteration of the GA, a set of members, also known as the population, is proposed. Each member is composed of a proposed value for each of the nine coefficients in question. During each iteration, every member in the population is evaluated and assigned a fitness value. The best members of the population are used to generate a new population for the next iteration. This process is known as crossover and mutation. Each member has genes which are the properties of the member used to create new members

during crossover. For the purposes of this use case, the genes are the individual coefficients in each set of nine.

Selection of data

Each recursion of the GA operates on a different subset of the inactivity data collected in Step 2. Each subset is made up of 10% of the total inactivity data (selected to reduce computation time). Each recursion is fed a different subset so as to avoid overfitting the solution to one specific set of inactivity points. Each 10% subset is made up of data from data selected from the value subsets described in Step 2, and a small amount of randomly selected data.

Evaluation and fitness

In order to evaluate the population, all inactivity points in the 10% subset are scaled using the coefficients that make up the member in question. Fitness is determined by comparing the results of evaluation to the 1G vector magnitude. The smaller the difference between the scaled vector magnitude and the reference of 1G, the higher a fitness is assigned to that member, with the best hypothetical fitness being a difference of 0. Members are then ordered from best to worst based on fitness.

Selection of the next generation

After evaluation and assignment of fitness, a new population must be selected for the next generation. This new population consists of the best member from the previous population, a group selected through crossover and mutation, and four members whose coefficients are randomly generated. These last four members minimize the risk of GA convergence to a local minimum.

GA Iteration and Recursion

Following the GA technique, each recursion produces a new evolved generation of coefficients that are more fit than the previous. After all generations are complete, a new recursion of the GA is started using the final population of the previous recursion as the initial population. Sixteen total recursions are performed in total in order to maximize the likelihood of a stable solution being found. Each recursion is run on a different 10% subset of the data to avoid overfitting the solution to a specific set of inactivity points.

Step 4: Verify Fix

The solutions can be verified by examining the fitness of members as the GA iterates. As the GA approaches a solution, the fitness should trend toward and begin to dither about a central value. If the fitness continues to trend in a specific direction, or if the standard deviation of the fitness values is high, then the GA has not settled on a solution. If fitness dithers about 0 with little variation, then the solution is valid.

Step 5: Fix data

Once all coefficients have been discovered and verified, they can be applied to the affected data files. To do so, Equations 4, 5, and 6 can be applied to the raw (uncalibrated) file assuming that each data point represents the measured variable (e.g. $X_{measured}$) in Equation 4.

Results

To verify that the three methods work, affected and unaffected units were exposed to the same accelerations simultaneously. A visual plot of the raw data as well as a histogram of the errors for the File Scan method (Figure 4) the Calibration File Application method (Figure 5) and the Factory Calibration Method (Figure 6) show that total errors have been reduced. The histogram of errors can be obtained by identifying the envelopes of two comparable data series. It is clear in the raw data plots that the data collected from an unaffected unit and the fixed data from the affected unit show similar values, and the difference between the uncorrected and corrected data is visually clear. The histograms illustrate how the error distribution more closely clusters around 0 after the data is fixed.

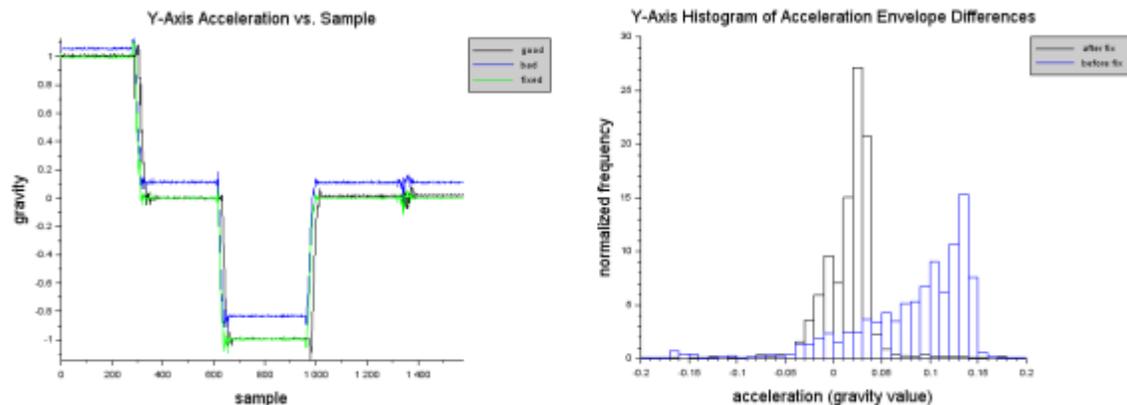


Figure 4 - File Scan method showing raw acceleration results and histogram of errors.

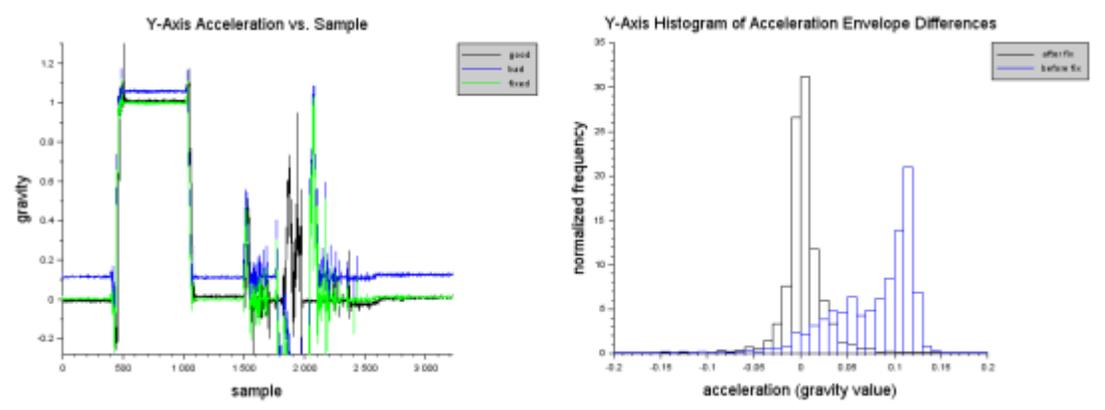


Figure 5 - Calibration File Application method showing raw acceleration results and histogram of errors.

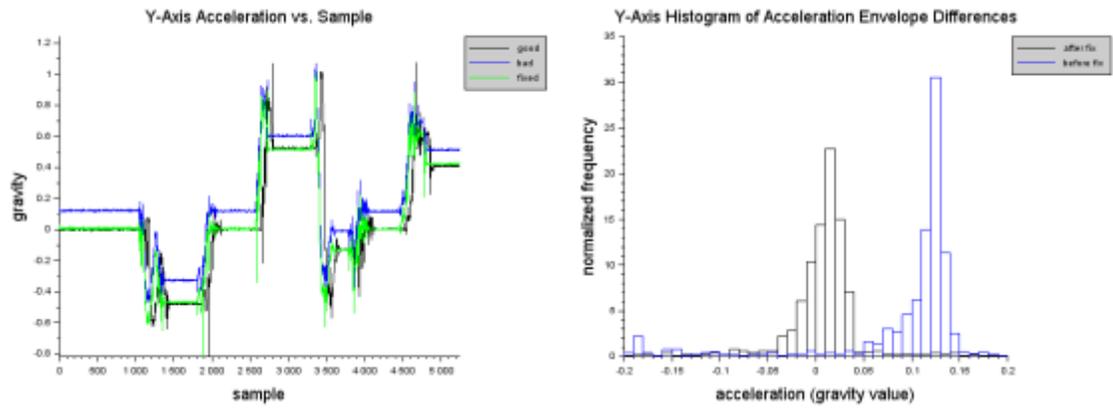


Figure 6 - Factory Calibration method showing raw acceleration results and histogram of errors.

Discussion

Combining the results of all three methods has led to a very high success rate when attempting to repair affected data files. In all cases where a unit was returned for factory recalibration, the affected data files were repairable. In almost all cases, where sufficient inactivity data was found, or a calibration file was provided, the GA was able to find a suitable set of coefficients.

In a small subset of cases, the GA can fail to find suitable coefficients. There are a couple of reasons why this might happen.

Local Minimums

The GA approach is, by nature, subject to the inherent issue of producing inaccurate results due to the convergence of errors into local minimums. Thus, there is no guarantee that a solution will be found, or that all possible solutions will be searched. The given approach minimizes this issue by creating diversity in the GA populations and running the GA recursively.

Temperature Drift

If the affected activity monitor was operating in a high ambient temperature environment, due to temperature drift. The reference data captured from the periods of inactivity may be skewed, thereby resulting in invalid coefficients following the GA process.

Software Tool

In order to expedite the file repair process, a software tool was created which combines each step of these processes into a single application. This tool automates the process of reading and scanning the data files, selecting inactivity periods, and running the GA to find the new coefficients. It then applies the new coefficients and generates corrected data files.

Declarations and Acknowledgements

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

Sample data supporting and demonstrating the techniques developed during this research is available on the ActiGraph support website. ActiGraph has published sets of example data corresponding to the three methods, available at this URL:

https://actigraph.desk.com/customer/portal/kb_article_attachments/94857/original.zip.

Funding

All funding for this project came from ActiGraph, LLC. ActiGraph provided device data on which the research was based.

Authors' contributions

The specific techniques described herein were developed by Joe Nguyen, with written and editorial support from Jeremy Wyatt and Nathan Miller.

Competing interests

All authors are employees of ActiGraph, LLC. This research was an internal project of ActiGraph, LLC. As such ActiGraph was the sole source of funding and research was carried out by ActiGraph employees. Results of this work do not constitute endorsement by the ACSM. The results presented herein are presented in their entirety without fabrication or inappropriate manipulation. The authors declare that they have no competing interests.

Conclusions

Three methods have been formulated to repair files affected by the 16CPAN03 issue. The first method requires only the affected data files and the 16CPAN03 File Scaler Tool. This method presents the least burden to customers with affected data. The second method, while similar, requires the original affected activity monitor be present (to obtain a baseline) and is thus slightly more burdensome on customers. Due to the inherent potential of the discovery of local minimums or simple failure to converge on the global minimum, it is possible that these first two methods may be unsuccessful at recovering the proper calibration coefficients. While more burdensome, the third method repeats the factory calibration procedure on the hardware, thereby guaranteeing successful recovery of impacted files. All three methods have empirically proven to be successful at recovering and reversing the 16CPAN03 issue for impacted activity monitors and/or .gt3x files. Customers can reliably apply any of these methods to affected files by utilizing the software repair tool.

References

1. ActiGraph Support [Internet]. Pensacola: ActiGraph; c2016 Performing Calibration File Generation Routine for 16CPAN03; 2016 June 3 [Cited 2016 July 26]. Available from <https://help.theactigraph.com/entries/108679463>
2. Frosio I, Pedersini F, Borghese NA. Autocalibration of MEMS Accelerometers. IEEE Transactions on Instrumentation and Measurement. 2008 October 21;58(6). 2034-2041.
3. Hornby G, Globus A, Linden D, Lohn J. Automated Antenna Design with Evolutionary Algorithms. AIAA Space 2006;19-21.
4. Micropower, 3-Axis, ± 2 g/ ± 4 g/ ± 8 g Digital Output MEMS Accelerometer. Analog Devices; 2016 August 30 [Cited 2016 September 13]. Available from <http://www.analog.com/media/en/technical-documentation/data-sheets/ADXL362.pdf>
5. Pedley M. High-Precision Calibration of a Three-Axis Accelerometer [Internet]. [place unknown]: Freescale Semiconductor; 2015 Oct [Cited 2016 July 28] 35 p. Available from: http://cache.freescale.com/files/sensors/doc/app_note/AN4399.pdf
6. Whitley D. A genetic algorithm tutorial. Stat Comput. 1994;4(2):65-85.

Figure captions

- Figure 1. Timeline showing the sequence of events.
- Figure 2. Z axis measured accelerations. Good unit readings cluster near -1 earth gravity. Bad units cluster between -0.8 and -0.85 earth gravity.
- Figure 3. A sample of inactivity data showing points in each value-subset, with the valid inactivity period appearing in the -0.7G to 0.7G subset.
- Figure 4. File Scan method showing raw acceleration results and histogram of errors.
- Figure 5. Calibration File Application method showing raw acceleration results and histogram of errors.
- Figure 6. Factory Calibration method showing raw acceleration results and histogram of errors.