Characterizing Phone Battery Charging Behaviour Using Sensor Data and Unsupervised Learning

Janelle Berscheid
University of Saskatchewan
jrb871@mail.usask.ca

Binben Wang
University of Saskatchewan
biw123@mail.usask.ca

ABSTRACT

Cell phone batteries are usually thrown away when their performance degrades, leading to an increase in waste entering landfills, and an increase of toxic chemicals entering the ecosystem and damaging human health. Since phone users' battery charging habits can impact battery life, if we can identify the charging behaviours of users, then this knowledge could be used to develop better batteries or smarter applications, leading to less battery waste in the future. In this project, we propose that phone users may be clustered based on their battery-charging habits using two algorithms: k-means clustering, and lexical analysis, using features derived from phone sensor data. Applying both methods verified the existence of two clear phone charging patterns: charging the phone overnight, and charging as often as possible throughout the day.

KEYWORDS

unsupervised learning, sensor data, battery, gps

1 INTRODUCTION

Electronic waste has become the fastest-growing source of solid waste globally [8], as cell phones and other small consumer electronics have become ubiquitous. Lithium-based batteries are widely used in these small electronics because of their larger energy density and longer life expectancy. Another advantage of Lithium-based batteries is that hey can be charged at any charge level without much harm to battery capacity [11]. However, Lithium-based batteries start aging immediately after manufacture, even if not used [3]. When the performance of these batteries degrade, most of them wind up in a landfill because of a lack of regulation surrounding their disposal, and a low participation of consumers in battery recycling programs. Toxic materials in lithium-based batteries, like chromium and lead, pollute the environment and damage human and animal health [6]. For example, in Guiyu, China, a common dumping site for electronic waste, the mean blood concentration of lead in children was 15.3 μg/dl, well above the 10 μg/dl threshold recommended for remedial action [8]. Lead poisoning can be fatal in high doses, and can cause developmental delay in children.

Cell phone batteries often wind up in a landfill when participants upgrade their phones–usually because their current battery's performance has degraded. However, the way that participants charge their phone batteries can have a significant impact on the life of the battery. For example, Ferreira et al.'s survey of participant's battery charging habits found that on average, participants keep the phones plugged for 4 hours and 39 minutes after charging has been completed, which leads to energy waste and disrupts the charging cycle of the battery, reducing battery life [5]. Common advice about how to handle batteries and extend their life has changed over the years, and old advice may no longer be valid. Samsung, the largest Android phone manufacturer, warns participants that the old advice of completely discharging a battery is no longer valid for newer types of smartphone batteries, and that battery charge should be kept above 20% to preserve battery life [1].

Researchers and designers of mobile devices cannot currently identify details of how participants charge their phone batteries, which leads to energy waste and reduced battery life. Understanding how, when and why people charge their phones could inform the design of future batteries, phones, or applications designed to extend the life of a phone, in order to reduce the number of batteries thrown away.

Our goal was to explore the research question: is it possible to cluster individuals by where, when and how often they charge their phone batteries? We chose to compare and contrast two possible approaches of clustering users: first using k-means clustering, a classical unsupervised learning technique, and second by applying lexical analysis to participants' battery charge levels and identifying which participants have similar charging strings. We hypothesized that we would see three types of charging behaviour emerge:

(1) people who charge their phones only at night while they sleep
(2) people who charge their phones only when the battery drops below 20%, at any location
(3) people who charge their phones as often as possible (e.g. whenever near an outlet)

In this paper, we prove that the hypothetical patterns of charging overnight and of charging as often as possible can be identified by both a k-means clustering algorithm and lexical analysis of charge patterns, contingent on the participant using their phone for at least half of the study period.

2 RELATED WORK

Compared to other hardware improvements, improvement of battery energy density is at its slowest pace since 1990 [9]. The iPhone XS, one of the most popular phones in 2018, only has a 2658 mAh battery capacity, and so users need to charge their iPhone XS almost every day. To combat this problem, phone users adopt various methods to increase their battery lifetime, and lot of efforts have been devoted to trying to use phone batteries more efficiently [4, 7, 12, 14]. Several studies have affirmed that charging behaviors of cell phone users have an impact on batteries' longevity [2, 11, 13].

Previous studies have attempted to characterize battery use and recharge behavior on cell phones and laptop computers by using a variety of methods, surveys [2, 13] and interviews [2] being the most common. While these methods may provide qualitative and quantitative data about battery usage, they are based on self-reporting, and thus are prone to biases, such as social desirability bias, where participants in a study answer in a way that they think the researcher expects them to answer or in a way that makes them appear "good," rather than answering in a way that reflects their true behaviour.

Other studies collect battery traces using automatic logging tools [2, 5]. However, these methods come with some limitations. Automated logging tools which collect battery traces don't collect other relevant information, such as GPS location or app usage. Also, phone displays of battery information can be vague and inaccurate, which causes participants to develop different and inaccurate mental models of their phone batteries [15]. Because of this, users' charging habits may not provide the best performance for their phones. Ahmad Rahmati et al. provided qualitative and quantitative evidence of this problem [11]. They defined the concept of human-battery interaction (HBI) [10, 11], as cell phone users have to deal with limited battery lifetime through a reciprocal process, and claimed that cell phone users can be divided into two types based on the evidence of field trials [11]: the users who regularly charge their phone, regardless of the charge level, and users who charge their phone based on charge level feedback from the battery interface [11].

3 METHODS

Our analysis was conducted on the Ethica SHED 10 data. For this project, we examined data from the *battery* and *gps* tables within Ethica.Though Ethica usually has one battery record per duty cycle, a transfer error occasionally causes multiple observations to appear in the battery table for a duty cycle. These observations contain identical information, except the *record_time* field is offset by a few seconds. We removed these duplicate observations by averaging records in the same duty cycle.

3.1 Filtering

Prior to filtering, the raw data contained 107 participants. We calculated an estimated percentage of battery records per participant, based on the assumption that a participant would have a maximum of 8640 battery records over the course of the study (based on the assumption of 30 days of study, with 1440 minutes in a day, and a duty cycle every 5 minutes). We then filtered data by the following process:

(1) Remove any participants who do not have both battery and GPS records available.
(2) Average battery records from the same duty cycle, in the event that multiple records existed for the same duty cycle.
(3) Remove all participants who have less than 50% of the total possible battery records throughout the study.
(4) Remove records from all tables which correspond to GPS records which fall outside the bounds of the Greater Saskatoon area
$(52.058367, -106.7649138128)$, $(52.214608, -106.52225318)$.
(5) Remove all GPS records with error of greater than 100 m.
(6) Remove participant 972, as this participant's number of battery records drops below 50% after linking the battery data with GPS data by timestamp.

After filtering, 41 valid participants remained.

Since we were attempting to measure regular battery charging behaviour, we chose to exclude those participants who had less than

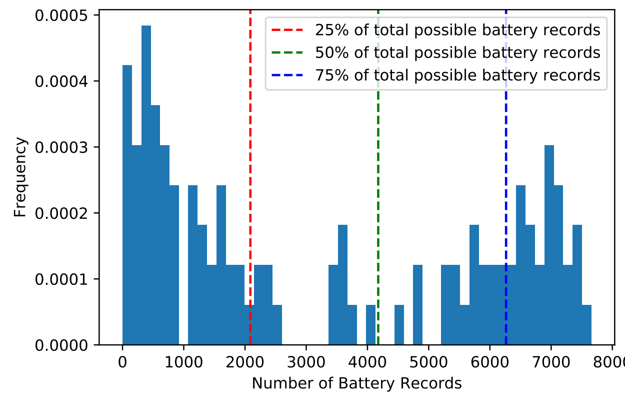

Figure 1: Number of Battery Records per Participant

50% of the possible battery table records available (see Figure 1). This is a reasonable threshold, since examining the battery data was our main goal, and fewer records may not give an accurate picture of participant behaviour.

While there are valid records outside of the Greater Saskatoon area, we wanted to capture a participant's ordinary behaviour, and traveling (which is infrequent among participants) may impact how a participant charges their battery. Therefore, we opted to exclude GPS points (and other data associated with those timestamps) which fell outside of the city limits. After filtering these points, participant 972 falls below the 50% battery records threshold, which suggests that this participant was traveling during many of the study days; this justifies their removal from the study.

GPS records with error greater than 100 m were eliminated, as the accuracy on these points exceeds the cell size of the grid when discretizing space in the following steps.

3.2 Stratifying and Aggregating

Once the data was filtered, the battery data was linked with the GPS data, by closest record time, to facilitate stratification across all types of features. The primary method of stratification is by whether the battery is charging or not, represented by a non-zero value or a 0 respectively in the *plugged* variable in the *battery* table. Battery records can then be further stratified by battery level while charging or discharging. The distribution of the battery levels in Figure 2 suggests that we should consider the battery level at 100% as a separate range, as it occurs at a much higher frequency than any other battery level value in the study. Therefore, we chose to bin battery level observations into three categories: 0-49%, 50-99%, and 100%. Another possible stratification of the battery charging records is by type of charging, such as being plugged into a wall or a USB port, as described in the *plugged* variable in the *battery* table.

Once the above battery stratifications were completed, the data was aggregated by the following process:

- Plug-in battery events were found by detecting changes from discharging to charging (or vice versa for unplugging events).
- Using these marked plug-in and unplugging events, charge intervals were grouped: all battery observations between

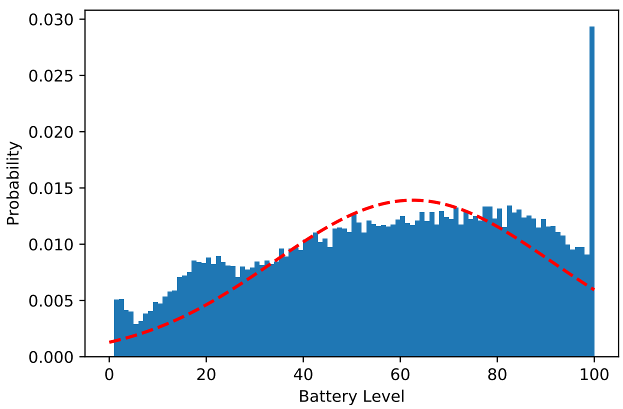

Figure 2: Battery Level Distribution

a plug-in and unplug battery events were considered to be from the same charge interval. These charge intervals could then be aggregated by their length or charge level.

- The city of Saskatoon was discretized into a grid with a cell size of 100 m × 100 m, which is small enough to capture differences between buildings on campus or between city blocks. This allowed the GPS data to be aggregated by mean location per duty cycle, or by charge/discharge interval.

3.3 Modeling

In the final stage of the data processing pipeline, we characterized battery behaviour by operationalizing the data into six categories of features:

- Number of plug-in events
- Average charge length (in duty cycles) per participant
- Percentage of time charging in different charge modes per participant
- Percentage of plug-in events at participant's home
- Number of unique locations charged in per participant (100m grid cells)
- Percentage of time (or duty cycles) in each of three battery level ranges (0-49, 50-99, 100)

These created a final feature vector with 10 features per participant.

3.3.1 Number of plug-in events per participant. This feature describes how many times the participant charged their phone over the course of the study. Its value is a positive integer, and is calculated from the aggregated data as follows:

(1) Stratify the data by participant.
(2) For each participant, count the number of plug-in events by detecting changes of the *plugged* variable from 0 (discharging) to 1 or 2 (charging).

3.3.2 Average charge length (in duty cycles) per participant. This feature describes the average length in duty cycles of each participant's charging intervals, rounded up to the nearest integer. Its value is a positive integer, and is calculated from the aggregated data as follows:

(1) Stratify the data by participant.
(2) For each participant, group the battery records into distinct charging intervals by when the *plugged* variable is equal to 0 (discharging) and is not equal to 0 (charging)
(3) Count the number of duty cycles in intervals in which the *plugged* variable is not equal to 0 to get the length of each interval.
(4) Take the average length of all of the charging intervals for each participant.

3.3.3 Percentage of time (dis)charging per participant. These three features describe percentages of time the participant was discharging their phone, charging their phones in mode 1 (charging from a wall socket) and charging their phones in mode 2 (charging from a USB port) over the course of the study. These features are represented by a vector in $[0, 1]^3$ whose components sum to 1, and is calculated from the aggregated data as follows:

(1) Stratify the data by participant.
(2) For each participant, count duty cycles in which the *plugged* variable is equal to 0, 1 and 2 and count total duty cycles.
(3) Divide the counts for each mode by the total number of duty cycles.

3.3.4 Percentage of plug-in events at participant's home. This feature describes the percentage of plug-in events that happened at the participant's home. A high number suggests the participant charges their phones at night regularly, while a low number suggests the participant charges their phone at any opportunity. Its value is a decimal in $[0, 1]$, and is calculated from the aggregated data as follows:

(1) Stratify the data by participant.
(2) For each participant, find the grid cell where the participant's home is located by finding the participant's most commonly inhabited grid cell in the time interval from 3 A.M. to 6 A.M..
(3) Filter battery records by the participant's *user_id* and the grid cell in which the participant's home is located.
(4) Count the number of plug-in events that happened at home by detecting changes of the *plugged* variable from 0 (discharging) to 1 or 2 (charging).
(5) Divide this count by the total number of plug-in events per participant.

3.3.5 Number of different grid cells charged in per participant. This feature is a count of the number of different grid cells the participant has charged their phone in at least once, with a grid size of 100 m. A low number means the participant charges their phone habitually in a small number of locations, while a high number suggests the participant charges their phone at any opportunity, regardless of location. Its value is an integer of 0 or greater, and is calculated from the aggregated data as follows:

(1) Calculate the grid cell index for each GPS point associated with a battery record.
(2) Stratify the records by charge status, and filter to get all records where the phone is charging.
(3) Stratify the data by participant.
(4) For each participant, count the number of unique grid cell indices associated with battery charging intervals.

3.3.6 Percentage of time in battery level ranges (0-49, 50-99, 100). This feature describes percentage of time battery levels of the participant's phone are in the ranges [0,49], [50,99] and [100]. A 100 battery charge level is more frequent than other battery levels, so we treat it as a separate range. These features are represented by a vector in $[0, 1]^3$ whose components sum to 1, and is calculated from the aggregated data as follows:

(1) Stratify the data by participant.
(2) For each participant, count the total number of duty cycles. Stratify battery data into battery level ranges [0,49], [50,99] and [100] and count the number of duty cycles in each battery level range.
(3) Divide the number of duty cycles in each battery level range by the total number of duty cycles.

These ten features, once calculated, enable the first clustering technique we applied: k-means clustering.

3.4 Approach 1: k-means Clustering

Our first approach involves unsupervised learning using the k-means algorithm, a common method of clustering data and finding patterns. Because of the curse of dimensionality, 10 features is too many to represent our dataset, since our filtered data has only 41 participants. Therefore, we first used principal component analysis (PCA) to reduce the number of features by selecting relevant features with the highest variance. We then adopted the elbow method to decide the optimal number of clusters and cluster our dataset by k-means. Finally, the detected clusters were compared using Silhouette scores (the mean Silhouette Coefficient over all samples), which compare the intra- and inter-cluster distances for each point, in order to determine good the clustering is by how far apart the final clusters are. We used the "scikit-learn" Python library (version 0.20.3) to apply these methods.

3.5 Approach 2: Lexical Analysis

Lexical analysis is a method of converting a sequence of characters into tokens and then examining the properties of that sequence of tokens. In this case, our goal was to compare sequences of participants' battery charge levels and quantify how similar they are, then cluster participants based on their level of similarity to each other. Applying lexical analysis to participants' battery data required only the first filtering step of the data pipeline, and then the following process was used to characterize the sequences of sensor data:

(1) Bin the battery charge level observations in 5% increments by dividing the charge levels by 5 and taking the floor of the result.
(2) For each user, convert the sequence of battery level readings to a string of integers separated by space characters, and then find the counts of unigrams and bigrams for the string.
(3) Calculate the TF-IDF (Term Frequency, Inverse Document Frequency) score for each vector.
(4) Calculate the pairwise cosine similarity between each participant's vector.
(5) Use the pairwise cosine similarity as an adjacency matrix and build a graph using values at or above the 75% and 90% quantiles of the cosine similarity.

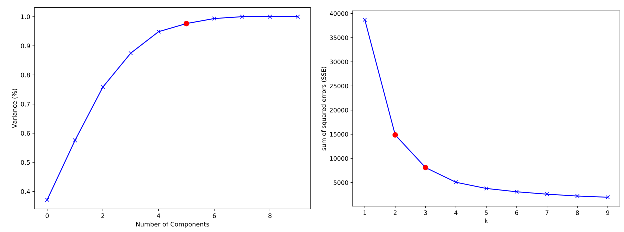

Figure 3: Principal Component Analysis Results

(6) Find the communities in the graph using Clauset-Newman-Moore greedy modularity maximization.

We chose to bin charge levels to 5% increments in order to reduce artificial differences between participants caused by the granularity of the data; for example, the difference between two pairs of observations "99, 98" and "99, 97" is not very significant on a scale of battery levels from 0 to 100, but the bigrams of these two strings would suggest they are different. Both would become "19, 19" after binning data in 5% increments, and so their cosine similarity would be higher.

Both unigrams and bigrams were chosen for the representation of the sequences because this allows the feature vector to capture both the raw distribution of counts in each battery level bin as well as the transitions between two different observations (increases, decreases, or remaining at the same level). TF-IDF scores were chosen because they give higher weight to rarer terms; this is important because the battery level data is dominated by 100% charge level observations, as seen in Figure 2.

Since the cosine similarity is calculated pairwise for all participants, the resulting graph will be fully connected if the edges are not filtered. In order to find only the most significant connections, we chose to filter at the 75% and 90% quantiles, to keep only the top 25% and 10% of connections. This allows us to compare how the clusters change under increasingly strict similarity constraints.

Steps 2, 3 and 4 of this process were all completed using the "scikit-learn" Python library (version 0.20.3), while Steps 5 and 6 used the "NetworkX" Python library (version 2.2).

4 RESULTS

For simplicity, we will use the following numbers to refer to different features.

1 Percentage of plug-in events at participant's home;
2 Percentage of time discharging;
3 Percentage of time charging in mode 1 (wall socket);
4 Percentage of time charging in mode 2 (USB charge);
5 Number of plug-in events;
6 Average charge length (in duty cycles);
7 Number of unique locations charged in per participant;
8 Percentage of time battery level between 0 and 49;
9 Percentage of time battery level between 50 and 99;
10 Percentage of time battery level is 100.

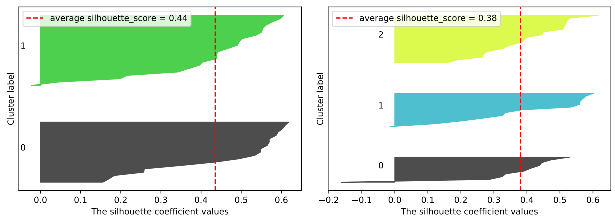

Figure 4: Silhouette Scores for clustering when k=2 and k=3

4.1 k-means Clustering

The left figure in Figure 3 depicts the relationship the between number of principal components and corresponding cumulative explained variance. We see that five principal components almost cover 95% of total variance, which reduces the number of features by half while maintaining most of the variance. The right figure in Figure 3 shows the results of applying the elbow method: the relationship between number of clusters and corresponding sum of squared errors (SSE). $k = 2$ and $k = 3$ are both good candidates for k-means clustering from the elbow method's perspective, so we clustered participants into two or three clusters and then compared the results of different clusterings.

Figure 4 visualizes the Silhouette scores of participants in the two- and three- cluster cases. In general, participants have higher silhouette scores in the two-cluster case: the average Silhouette score is 0.44, which is slightly better than 0.38, the average Silhouette score in the three-cluster case. Noticeably, some participants in the three-clusters case even have negative Silhouette scores, which indicates that participants in that cluster are not similar enough to each other.

Figure 5 shows boxplots of all 10 features in the two-cluster case. Features 1 (percentage of plug-in events at participants' home), 5 (number of plug-in events) and 6 (average charge length) have significant differences between clusters. Figure 6 shows boxplots of all features in the three-cluster case. This shows that the same three features still have significant differences between clusters in this case; additionally, Features 2 (percentage of time discharging) and 3 (percentage of time charging in mode 1) also show significant differences.

4.2 Lexical Analysis

Applying lexical analysis and limiting the results to the 75% and 90% quantile results in two and three major clusters respectively. As seen in Figures 7 and 8, where nodes are coloured based on their assigned community, there are clusters of nodes which are similar to each other and dissimilar to members of other clusters, with participant 1332 serving as a bridge between these groups. (Note that nodes with no edges above the quantile threshold are not displayed in these graphs, and are considered to belong to no cluster.)

To evaluate the clusters found by lexical analysis, we can compare the mean and standard deviation of the features created in the data processing phase. The results in Tables 1 and 2 show that most of the features within these groups of clusters have no significant difference between them. In particular, the column for feature 5 (number of charge events, highlighted in red) shows no difference, even though this was the most significant feature found in the k-means clusters. However, the feature 10 (percentage of battery levels at 100%), which showed no significant difference in the k-means clustering, is the strongest differentiator between the lexical analysis clusters. This makes intuitive sense, as the vectors calculated for each participant capture both the number of duty cycles spent at 100% (unigrams) and the number of consecutive duty cycles where a participant remained at 100% (bigrams); on the other hand, information like GPS data or the charging mode is not captured by the lexical analysis, and so is not a differentiating feature in the resulting clusters.

Table 3 summarizes the agreements between k-means and lexical analysis, which is the intersection between the clusters created by both methods. k-means and lexical analysis do show some agreement: when considering 2 clusters, they agree on the clustering of 21 out of 41 participants, but they only agree on clustering of 12 participants when considering 3 clusters. The two distance measures capture different facets of the data, and so do not completely agree on which participants are close together and which are far away. This is demonstrated in Figure 9 (coloured by k-means cluster assignment), which compares the normalized pairwise Euclidean distance of principle components with the normalized 1- cosine similarity values between each participant. If these measures agreed on which participants were close together and which were far away, this plot would show a roughly linear relationship between the two methods; however, the two methods have almost no agreement on which two points are far away from each other, regardless of the cluster.

4.3 Comparison between k-means and Lexical Analysis

Figure 10 shows a visualization of the battery level records of participants that k-means and lexical analysis both agree on. In each sub-figure, we plot two participants' time series. The plots in the left column visualize when k-means and lexical analysis both create two clusters, while the plots in the right column visualize when k-means and lexical analysis both create three major clusters. One clear pattern emerges in both cases: participants who regularly charge their phone at night (seen in the bottom left and top right plots). In these cases, participants discharge throughout the day and charge to full at night.

Though the graphs in the top left and middle right seem more chaotic, this in itself suggests a pattern of charging as frequently as possible, regardless of time or location. These users' phone batteries charge and discharge rapidly, and it is rarer that a charge interval brings the phone up to 100%, or that the phone drops close to 10%. The bottom-right plot does not show an obvious pattern. The participants here seem to sometimes charge their phones at night, sometimes not, but the blue line indicates a participant who is not a heavy smartphone user, taking a long time to deplete their battery. This may be the cluster that "atypical" users get assigned to.

These results show that two of our three hypothesized charging patterns–charging at night and charging as often as possible–can be detected by both clustering methods.

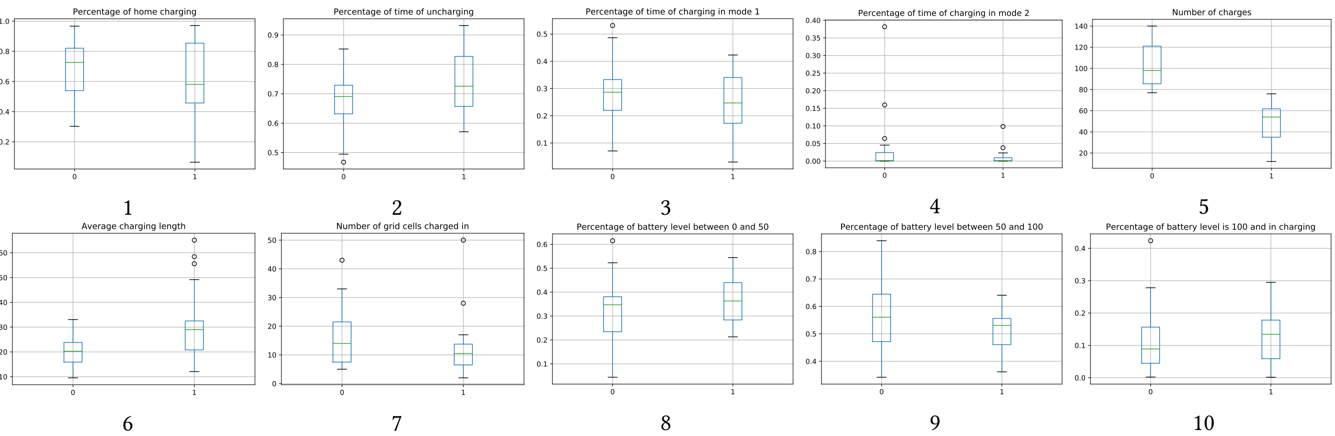

Figure 5: Boxplots of Features in 2-means

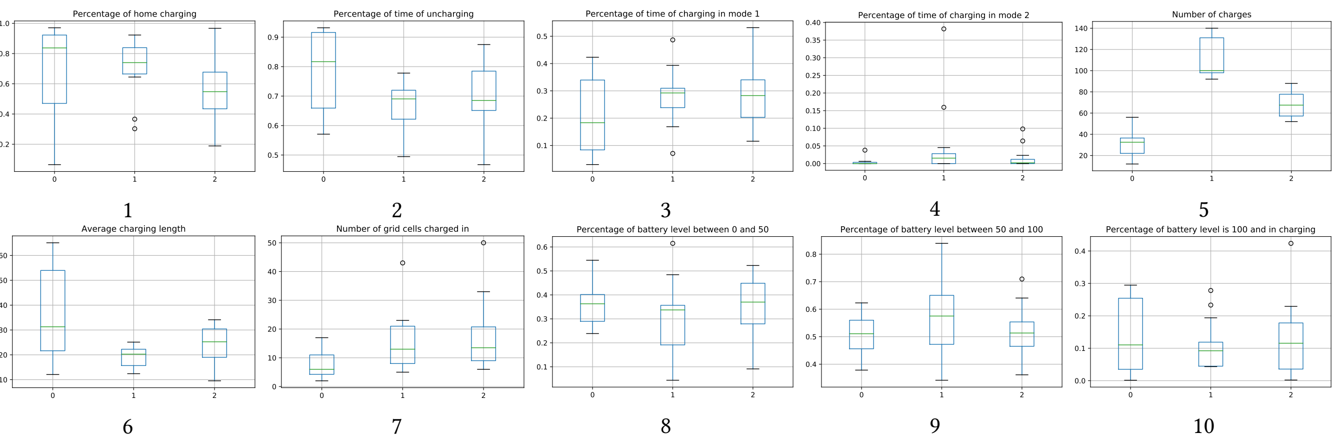

Figure 6: Boxplots of Features in 3-means

Table 1: Means of Features Lexical Analysis Clusters

Features		1	2	3	4	5	6	7	8	9	10	cnt
75% Quantile	0	0.67	0.65	0.32	0.03	72.48	31.87	15.05	0.27	0.53	0.20	21
	1	0.63	0.78	0.20	0.02	76.22	19.33	12.72	0.40	0.56	0.05	18
90% Quantile	0	0.71	0.74	0.24	0.02	84.38	21.06	13.62	0.41	0.53	0.06	13
	1	0.59	0.68	0.28	0.04	73.20	30.25	12.7	0.26	0.57	0.16	10
	2	0.76	0.58	0.42	0.004	76.29	40.03	16.71	0.23	0.48	0.29	7

Table 2: Standard Deviation of Features in Lexical Analysis Clusters

Features		1	2	3	4	5	6	7	8	9	10	cnt
75% Quantile	0	0.22	0.08	0.10	0.08	31.79	13.80	12.73	0.13	0.13	0.08	21
	1	0.26	0.09	0.08	0.04	36.59	6.28	7.31	0.07	0.07	0.03	18
90% Quantile	0	0.22	0.08	0.07	0.05	37.71	6.18	8.13	0.07	0.06	0.02	13
	1	0.17	0.07	0.09	0.12	10.26	6.18	11.31	0.06	0.09	0.04	10
	2	0.22	0.07	0.07	0.006	37.08	17.35	9.95	0.11	0.08	0.07	7

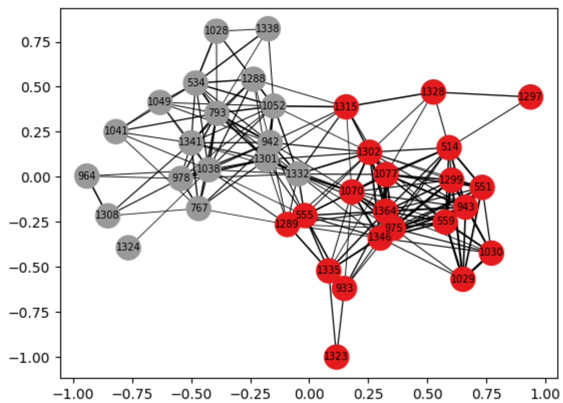

Figure 7: Communities in 75% quantile for cosine similarity

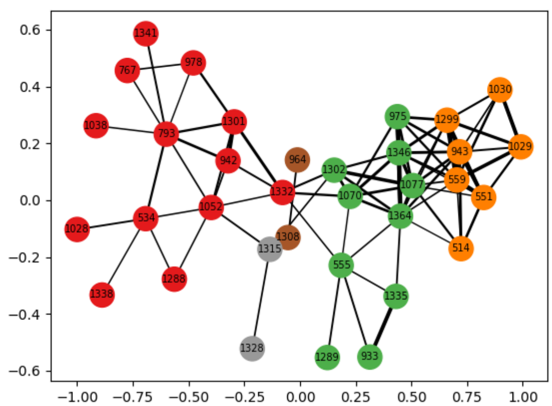

Figure 8: Communities in 90% quantile for cosine similarity

5 DISCUSSION AND SUMMARY

Since both k-means and lexical analysis were able to identify two clear patterns in the data–participants who charge at night and participants who charge at every opportunity–even though they both measured different aspects of the data, this suggests that these patterns could be upheld by further clustering using other unsupervised learning methods. The results of applying these clustering

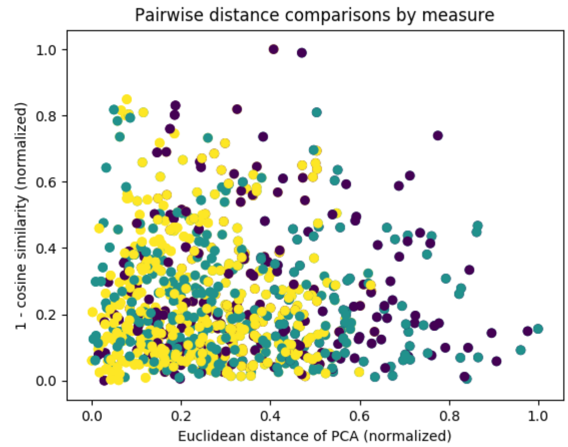

Figure 9: Comparison of pairwise distance measures between clustering methods

methods could generalize to other cities and other studies; features that determined the results of the clustering had little to do with the city of Saskatoon, even though we filtered data to exclude non-Saskatoon data points.

Table 3: Agreements between k-means and Lexical Analysis

		# of Agreements	Total
2-means	0	10	21
	1	11	
3-means	0	3	12
	1	6	
	2	3	

5.1 Limitations

The main limitation of this study is that collecting the data itself had an impact on all participants' battery discharging behaviour: the Ethica app drains phone batteries by collecting and transmitting sensor data at regular intervals. This would disrupt participants' regular phone usage patterns, and could have caused them to charge more frequently than usual. Another limitation is that some of the data was collected during the Saskatoon winter or early spring, when cold temperatures outside could have affected the performance of the battery or the readings taken by the battery sensors. It was not possible to correct for these limitations in the data collection.

5.2 Future Work

We are interested in further exploring the characteristics of users who do not charge regularly at night. There may be aspects of these participants' behaviours which are not adequately captured by the ten features we identified; visualizations of their battery or GPS data may reveal clues about where to find more information about their habits.

Additional features may come from other sources of data within the study. When operationalizing battery charging characteristics of participants, we only considered data from the *battery* and *gps* tables only, but future work may include data from other Ethica tables, especially *app_usage* and *bluetooth*. Both of these tables can give an indication of how heavily the phone is being used and how that may affect the battery discharging, but we did not incorporate this information into our analysis.

Lexical analysis applied to the raw battery level data for each participant identified the night-charging and frequent charging

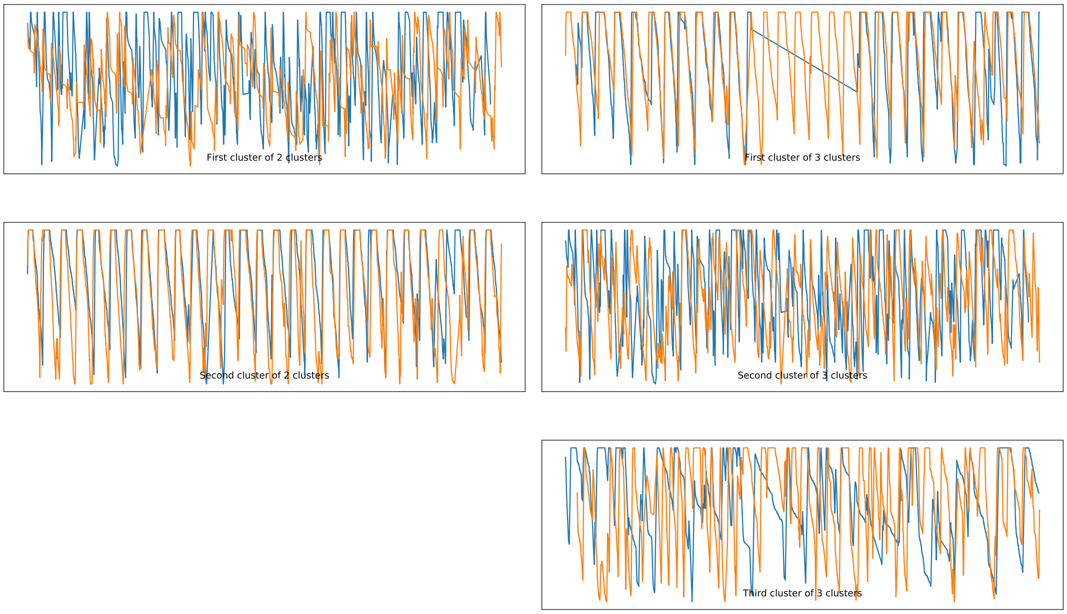

Figure 10: Visualization of Clusters in 2-means and 3-means

patterns, but a future approach may try applying the technique to a different facet of the battery data. For example, rather than analyzing a string of battery level observations over time, charging behaviour may instead by characterized as a string in some other way–for example, as a series of charge deltas.

5.3 Class Work Summary

Over the course of this project, we were able to gain more experience with clustering and unsupervised learning, which neither of us had worked with on a real project before. It helped highlight the tradeoffs between different parameters, as well as showed how difficult it can be to draw conclusions from your results without a ground truth to compare to, and what good metrics for comparing clusters might be in different scenarios. Learning about Luana's research with behavioural data and lexical analysis gave us a unique opportunity to get hands-on experience with an unusual technique, and to evaluate how it might be appropriate for other kinds of problems we may encounter in the future.

Given the chance to begin the project again, we would have liked to incorporate other tables of data, like *bluetooth* and *app_usage*, to characterize participants' charging behaviours in different ways. We also would have started incorporating the lexical analysis into the project at an earlier stage, rather than as an afterthought, to give us a better chance to explore the different possibilities of applying it and to come up with better visualizations of its results.

6 CONCLUSION

By applying k-means clustering and lexical analysis to phone usage data, we were able to identify two clear patterns of charging behaviour in participants: charging their phones at night, and charging their phones as often as possible. Both methods affirmed these patterns, which suggests that they can be observed using other methods. This gives researchers additional tools for understanding and being able to identify users' charging patterns from their phone usage data. In the future, this could give designers insight into how to build better, more eco-friendly batteries, or how to prompt users to adopt behaviours that will preserve the life of their phone batteries.

REFERENCES

[1] How can I optimise and extend the battery life on my Samsung Galaxy smartphone? | Samsung Support UK. http://www.samsung.com/uk/support/mobile-devices/how-can-i-optimise-and-extend-the-battery-life-on-my-samsung-galaxy-smartphone/. Accessed: 2019-03-10.

[2] Nilanjan Banerjee, Ahmad Rahmati, Mark D. Corner, Sami Rollins, and Lin Zhong. Users and Batteries: Interactions and Adaptive Energy Management in Mobile Systems. In John Krumm, Gregory D. Abowd, Aruna Seneviratne, and Thomas Strang, editors, *UbiComp 2007: Ubiquitous Computing*, volume 4717, pages 217–234. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

[3] Isidor Buchmann et al. *Batteries in a portable world: a handbook on rechargeable batteries for non-engineers*. Cadex Electronics Richmond, 2001.

[4] Soumya Kanti Datta, Christian Bonnet, and Navid Nikaein. Android power management: Current and future trends. In *2012 The First IEEE Workshop on Enabling Technologies for Smartphone and Internet of Things (ETSIoT)*, pages 48–53. IEEE, 2012.

[5] Denzil Ferreira, Anind K. Dey, and Vassilis Kostakos. Understanding Human-Smartphone Concerns: A Study of Battery Life. In Kent Lyons, Jeffrey Hightower, and Elaine M. Huang, editors, *Pervasive Computing*, volume 6696, pages 19–33. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[6] Daniel Hsing Po Kang, Mengjun Chen, and Oladele A. Ogunseitan. Potential Environmental and Human Health Impacts of Rechargeable Lithium Batteries in Electronic Waste. *Environmental Science & Technology*, 47(10):5495–5503, May 2013.

[7] Amilcar Meneses-Viveros, Erika Hernández-Rubio, Sonia Mendoza, José Rodríguez, and Ana Belem Márquez Quintos. Energy saving strategies in the design of mobile device applications. *Sustainable Computing: Informatics and Systems*,

19:86–95, 2018.

[8] O. A. Ogunseitan, J. M. Schoenung, J.-D. M. Saphores, and A. A. Shapiro. The Electronics Revolution: From E-Wonderland to E-Wasteland. *Science*, 326(5953):670–671, October 2009.

[9] Joseph A Paradiso and Thad Starner. Energy scavenging for mobile and wireless electronics. *IEEE Pervasive computing*, (1):18–27, 2005.

[10] Ahmad Rahmati, Angela Qian, and Lin Zhong. Understanding human-battery interaction on mobile phones. In *Proceedings of the 9th international conference on Human computer interaction with mobile devices and services - MobileHCI '07*, pages 265–272, Singapore, 2007. ACM Press.

[11] Ahmad Rahmati and Lin Zhong. Human–battery interaction on mobile phones. *Pervasive and Mobile Computing*, 5(5):465–477, 2009.

[12] Nishkam Ravi, James Scott, Lu Han, and Liviu Iftode. Context-aware battery management for mobile phones. In *2008 Sixth Annual IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 224–233. IEEE, 2008.

[13] Saurabh Saxena, Gustavo Sanchez, and Michael Pecht. Batteries in Portable Electronic Devices: A User's Perspective. *IEEE Industrial Electronics Magazine*, 11(2):35–44, June 2017.

[14] Kazuhiko Takeno, Masahiro Ichimura, Kazuo Takano, and J Yamaki. Methods of energy conservation and management for commercial li-ion battery packs of mobile phones. In *The 25th International Telecommunications Energy Conference, 2003. INTELEC'03.*, pages 310–316. IEEE, 2003.

[15] Khai N. Truong, Julie A. Kientz, Timothy Sohn, Alyssa Rosenzweig, Amanda Fonville, and Tim Smith. The design and evaluation of a task-centered battery interface. In *Proceedings of the 12th ACM international conference on Ubiquitous computing - Ubicomp '10*, page 341, Copenhagen, Denmark, 2010. ACM Press.