Boosting Service Availability for Base Stations of Cellular Networks by Event-driven Battery Profiling

Xiaoyi Fan
School of Computing Science
Simon Fraser University
Burnaby, BC, Canada
xiaoyif@sfu.ca

Feng Wang
Department of Computer and Information Science
The University of Mississippi
University, MS, USA
fwang@cs.olemiss.edu

Jiangchuan Liu
School of Computing Science
Simon Fraser University
Burnaby, BC, Canada
jcliu@cs.sfu.ca

ABSTRACT

The 3G/4G cellular networks as well as the emerging 5G have led to an explosive growth on mobile services across the global markets. Massive base stations have been deployed to satisfy the demands on service quality and coverage, and their quantity is only growing in the foreseeable future. Given the many more base stations deployed in remote rural areas, maintenance for high service availability becomes quite challenging. In particular, they can suffer from frequent power outages. After such disasters as hurricanes or snow storms, power recovery can often take several days or even weeks, during which a backup battery becomes the only power source. Although power outage is rare in metropolitan areas, backup batteries are still necessary for base stations as any service interruption there can cause unafforable losses. Given that the backup battery group installed on a base station is usually the only power source during power outages, the working condition of the battery group therefore has a critical impact on the service availability of a base station. In this paper, we conduct a systematical analysis on a real world dataset collected from the battery groups installed on the base stations of China Mobile Ltd co., and we propose an event-driven battery profiling approach to precisely extract the features that cause the working condition degradation of the battery group. We formulate the prediction models for both battery voltage and lifetime and propose a series of solutions to yield accurate outputs. By real world trace-driven evaluations, we demonstrate that our approach can boost the cellular network service availability with an improvement of up to 18.09%.

1. INTRODUCTION

The deep penetration of cellular mobile wireless technologies have made anytime anywhere communications become more available than ever. Indeed, with the wide deployment of 4G network as well as the emerging of 5G, mobile applications are experiencing an explosion while the bandwidth of cellular network becomes higher than ever. To afford such great demands, in the cellular network infrastructure, more and more base stations have been strategically constructed and deployed to satisfy the service coverage and quality (e.g., bandwidth) requirements. The base stations are then aggregated into the carrier network through the backhaul infrastructure by fiber lines or microwave links. However, this renders that the network service downtime will cascade to all the dependent base stations if any part experiences power outages [1]. In particular, in the rural areas or during severe weather conditions such as hurricanes or snow storms, the power recovery may take several days or even weeks depending on the difficulty to reach the repairing area.

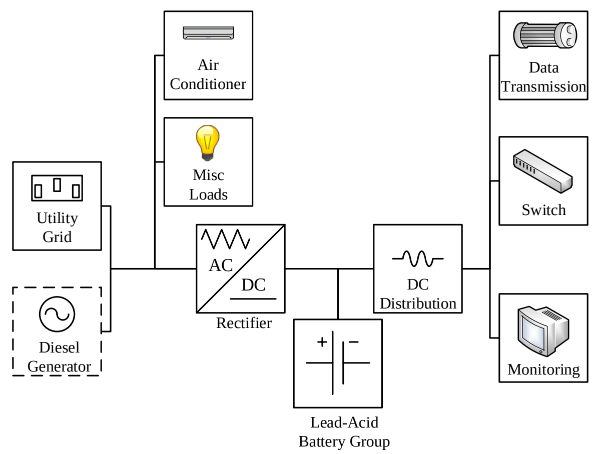


Figure 1: The power sources and loads in the cellular network base stations

To improve the service availability, besides connecting to the utility grids, each cellular tower is also equipped with a battery group as illustrated in Fig. 1. When a power outage happens in the utility grid, to avoid any service interruption, the battery group discharges to support the communication equipment until the generator is delivered to provide enough power supply. This makes the battery group and its working condition play a critical role. During this period, if the condition of the battery group deteriorates, a power outage can easily take down the service. The emergency repairing service arrives may take several days or even weeks depending on the difficulty to reach and locate the repairing section, especially in rural areas or during severe weather. Thus, being able to predict the condition of the battery group is of immense technical and commercial importance for system maintenance, so that battery replacement can be planned ahead of failure to minimize the interruptions of service availability. A successful prediction requires not only the knowledge of the battery ageing processes, but also a good understanding of the stress events that may induce and accelerate the aging process.

To tackle those issues, we collaborate with China Mobile Ltd co., which is the world's largest mobile phone operator in term of subscriber and currently with about 806 million users. We have collected 42,527 equipment data with totally 1,313,533,557 rows from July 28th, 2014 to July 31st, 2015 including 105 categories of events, which are further processed and analyzed by our cloud computing platform with Hadoop 1.2.1 and Hive 1.2.1. Based on the data analysis, we propose an event-driven battery profiling approach to precisely predict the battery group working conditions and schedule maintenance accordingly, so as to guarantee the high service availability for the cellular networks. We formulate the prediction models for both battery voltage and lifetime and propose a series of solutions to yield accurate outputs. By real world trace-driven evaluations, we demonstrate that our approach can precisely predict the battery voltage and lifetime with the RMS error no more than 0.01 v, which can further the improves the service availability of the cellular towers up to 18.09%.

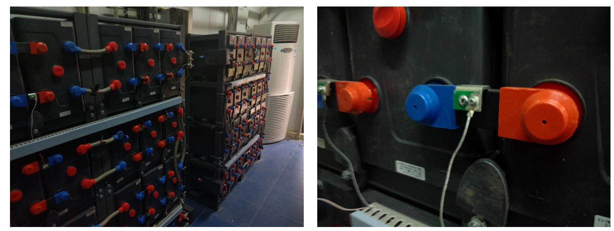
The rest of the paper is organized as follows. Section 2 provides our detailed observations on the battery properties. Section 3 presents our approach to efficiently solve the problem. Section 4 discusses the performance evaluation results on our approach. We provide a literature review in Section 5 and we conclude this paper in Section 6.

2. MOTIVATION AND DATA ANALYSIS

To better understand the battery working condition and its deterioration process, we have closely worked with the engineers from China Mobile to collect the dataset for our study. In this section, we first describe the collected dataset and then discuss our observations for aging profiling in time series battery data.

2.1 Background

We explain the details of power supply in the cellular network base stations, as Fig. 1 shows. The equipment in base stations is supported by the utility grid, where the battery group is installed as the backup power. In case that the utility grid interrupts, the battery discharges to support the communication switching equipment during the period of the power outage. As Fig. 2(a) shows, a battery group in the cellular network base station contains 24 cell batteries. In Fig. 2(b), the monitoring system has sensors on each cell of the battery group and periodically collects the voltage and events in both normal and abnormal situations. When the monitoring system reports the alert event, e.g. the power outage, the emergency repairing service is scheduled depending on the accident severity. Since few base stations have the diesel generators permanently installed on site, the engineers have to drive the emergency diesel generator to provide the power, which can take time to arrive at the site. The power outage can occur frequently and severely in the rural areas and developing countries due to the unstable utility grid. To make it even worse, the construction of infrastructure often makes that the base stations are difficult to reach, e.g. slippery rock trails in the mountains, where the workers have to manually carry the heavy generators to the site. As a result, the cellular carriers may have to trade off between the cost and the quality of service, so that they even abandon the base stations in the tough surroundings until the utility grid is restored. On the other hand, considering the labor costs for the mobile telecom carries, the periodical maintenance for the battery group is generally of long intervals, which further exaggerates the possibilities of battery accidents during the outage of utility grid. If the repairing engineers cannot arrive at the site before the battery group is exhausted, the availability of the base stations cannot be guaranteed. Thus prediction for the lifetime of battery group is meaningful for the service availability, which is helpful for maintenance engineers to solve the potential issues in advance during the periodical maintenance.



(a) Two battery groups with 24 cells

(b) The cable of monitoring systems

Figure 2: The battery group and monitoring systems

Despite the Li-ion and NiCd batteries demonstrate the latest development in battery technology due to their smaller size, lower weight and better storage efficiency, major drawbacks of these types of batteries are the high cost. Lead-acid batteries in Fig. 2(a) have large capacities and thus have been widely used for storage in backup power supplies in base stations. The aging mechanism of Li-ion batteries attracts many efforts [2], where the frequent activities of Li-ion batteries produce lots of log and provide possibilities to measure the battery working conditions. Yet the lead-acid batteries in base stations normally keep in the float-charging status, where float-charging status represents that a battery maintains the capacity by compensating for self-discharge after being fully charged. The monitoring system collects the float voltage from the float-charging batteries twice per day, which makes the dataset considerately sparse. Therefore extracting the features from such a sparse data source and predicting the working conditions of lead acid batteries pose many challenges and here we take the first attempt to tackle these issues.

2.2 Data Analysis

The log data that we have collected is from July 28th, 2014 to July 31st, 2015. In our dataset, we have identified 105 categories of event codes and obtained 42,527 equipment data with totally 531 tables and 1,313,533,557 rows.

2.2.1 *Voltage Readings of Batteries*

The voltage of each cell battery is the most important feature that we have measured, as it reflects the power output pattern of the battery. In general, we have observed two representative kinds of cell batteries, where we manually choose 1578 batteries as the newly-installed group and put 1459 batteries into the nearly-dead group depending on the repair records. As mentioned, there are 24 cell batteries in one battery group, where the rated voltage of cell is around 2.23v and the rated voltage of battery group is 53.5v. Based on this, we further analyze the typical status of the

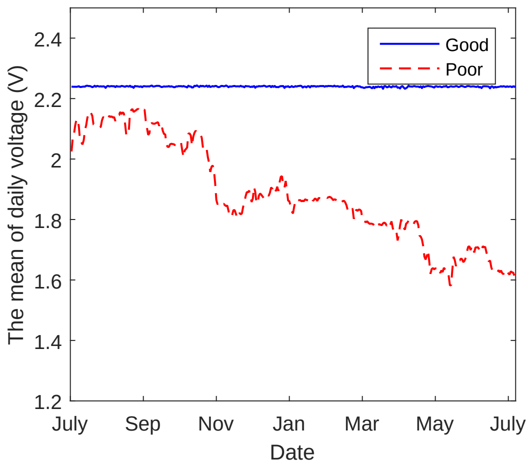


Figure 3: Mean voltage versus battery status

voltage patterns inside the two representative cell battery categories. Fig. 3 shows the significant differences between the newly-installed and nearly-dead batteries.The blue solid line plots the mean voltage of newly-installed batteries, which judders between 2.21v and 2.25v. The red dotted line shows the decay trend on the mean voltage of the nearly-dead batteries. There is a clear downward trend close to the failure date, where the battery power frequently falls down and becomes quickly exhausted, causing many issues and alerts in the cellular network base station, which indicates that the voltage have strong correlations with the battery life.

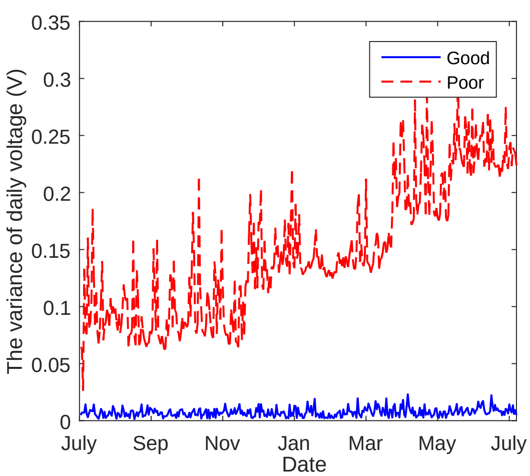


Figure 4: Voltage variances versus battery status

As Fig. 4 shows, the blue solid line represents the newly-installed battery can output a steady power and the variance of the voltage keeps very close to zero. The red dotted line illustrates that the variance of the nearly-dead batteries increases much faster than the newly-installed batteries. Fig. 4 illustrates that the voltage variance also has the correlation with the length of remaining lifetime. The variance of the output voltage from the batteries over time also reflects the aging trend of battery quality degradation. These observations motivate our battery working condition prediction based on the battery historical voltages.

2.2.2 *Battery Events*

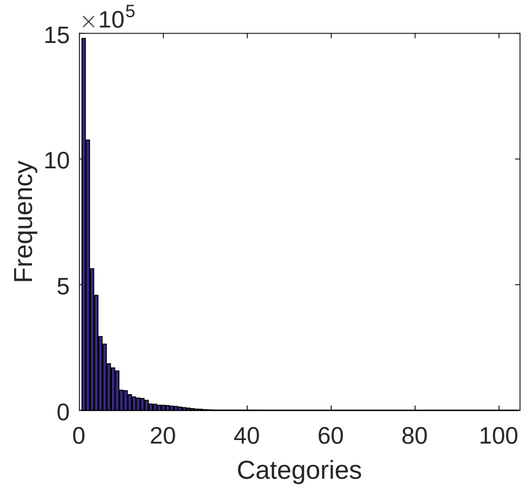


Figure 5: Distribution of battery events categories

We perform an analysis on the battery events in the logs to explore their potential relationship with the battery working conditions. Fig. 5 also shows the frequency distribution among all the 105 categories. We can see that the distribution is highly skewed: the most popular category is Alert, at about 28.09%, the second is Battery premature failure, at about 20.42%; and the third is Discharging, at about 10.70%.

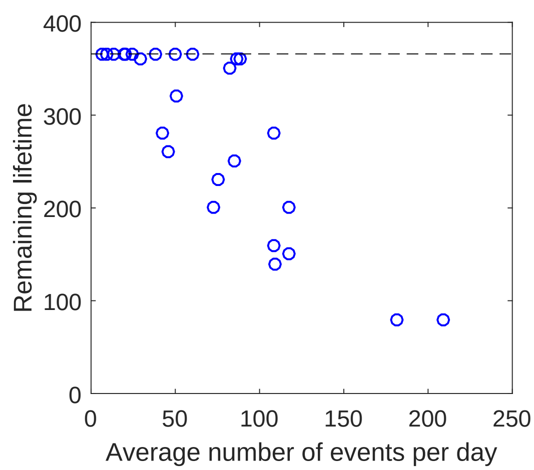


Figure 6: Correlation between the remaining life and the number of low float voltage events

We take three events as example to further investigate the correlations between events and battery remaining lifetime, which are *Alert to low float voltage*, *Discharge* and *fault cell* shown in Fig. 6, 7 and 8, respectively. We count the specific events number for each battery until the batteries are re-

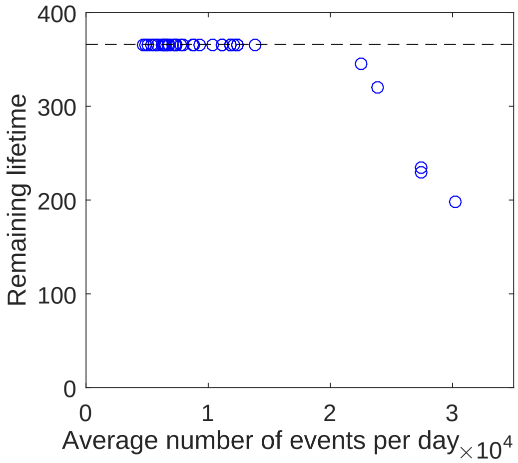


Figure 7: Correlation between the remaining life and the number of discharge events

placed, and pick up the top-30 batteries with the maximum number of events. From July 28th, 2014 to the end of July 2015, there are 366 days in our dataset, where the remaining lifetime of most batteries in our dataset is longer than 366 days, therefore dash lines represent that those batteries on it have longer remaining lifetime than 366 days. Fig. 6 and 7 plot the correlation between *Alert to low float voltage*, *Discharge* and remaining lifetime. This clearly demonstrates that there exists a strong correlation between battery remaining lifetime and *Alert to low float voltage*, as well as between battery remaining lifetime and *Discharge*. We further plot the failure rates against the number of *fault cell* events in the system in Fig. 8, which does not show any noticeable correlation between them. This implies that the failure is affected by some specific events. The observations suggest that the diverse events have different influences on the battery working conditions, thus it is necessary to discriminatingly differentiate these events for the accurate prediction.

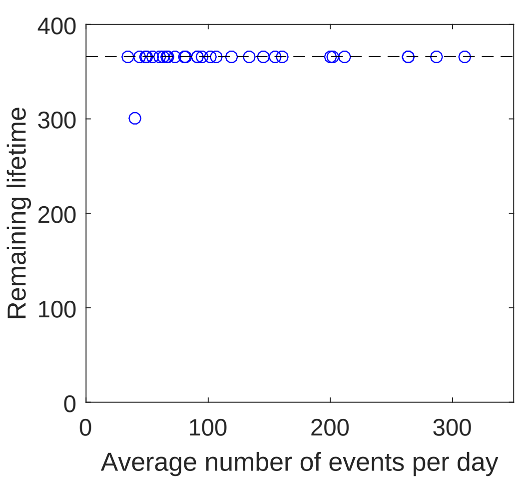


Figure 8: Correlation between the remaining life and the number of fault cell events

3. SOLUTION DESCRIPTION

The battery working conditions can be predicted by its relevant historical voltage records and event logs, which motivates us and serves as the basis for the prediction of battery remaining lifetime. Although the monitoring systems collects sufficient messages from batteries with the voltage and events data, there still remain several challenges to accurately predict the battery lifetime, especially when there are massive events in logs with some noises existing. Moreover, our observations suggest that the historical events are correlated with the battery working conditions, yet single event record is not a reliable factor for the prediction. Thus our approach filters out a large amount of noise and only pulls the relevant data from the database, including the historical float voltage and meaningful events, as mentioned in prior section, e.g., we only select the float voltage from the table named *historybattery*. Let $\mathcal{V} = \{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_N}\}$ denote the float voltage set and $\mathcal{E} = \{\mathbf{e_1}, \mathbf{e_2}, ..., \mathbf{e_N}\}$ represent the event set, where N is the number of batteries. Each voltage records $\mathbf{v_i} = \{v_i^1, v_i^2, ..., v_i^t, ...\}$ is a sequence of voltage records in time series for battery i, where each v_i^t is the mean voltage value of battery i in time t and the time interval is one day. Similarly, let $\mathbf{e_i} = \{e_i^1, e_i^2, ..., e_i^{m_i}\}$ represent the event set of battery i, which has m_i events. Our goal is to profile the aging trends for the battery working conditions, and predict the battery remaining lifetime. Given the battery data with historical events $\mathcal{E}$ and float voltage $\mathcal{V}$ in time series, we can predict the future voltage and then estimate the battery remaining lifetime.

To achieve this, we extract the aging trend $\mathbf{v_i}_a$ from a given time series $\mathbf{v_i}$. We partition the voltage records $\mathbf{v_i}$ into K segments, and use the polynomial regression to fit each segment, where each voltage segment k having the initial voltage v_{ia}^k and the slope of line with value s_i^k to represent the voltage aging trend. To ensure that the trend $\mathbf{v_i}_a$ is monotonic, we can simply write the constraint as $v_{ia}^k \geq v_{ia}^{k+1}$. Given the aging trend $\mathbf{v_i}_a = \{(v_{ia}^1, s_i^1), (v_{ia}^2, s_i^2), ...\}$, we propose a concept of *penalty* p_i^k to represent the influences of a group of events on the voltage aging trend, which can be defined as $p_i^k = s_i^k - s_i^{k-1}$. Then we utilize the learning algorithm with events data $\mathcal{E}$ to predict the penalty value p_i on the voltage aging trend $\mathbf{v_i}_a$. Based on the predictive voltage in coming months, we can estimate the battery remaining lifetime. To this end, we further transform the problem into a Multi-Instance Multi-labels Learning [3] problem.

Let $\mathcal{E}$ denote the instance space for the historical events, and $\mathcal{P}$ define the set of penalty labels. We denote the training data by $\{(\mathbf{e_1}, \mathbf{p_1}), (\mathbf{e_2}, \mathbf{p_2}), ..., (\mathbf{e_N}, \mathbf{p_N})\}$ that consist of N examples, where a set of events $\mathbf{e_i}$ is called a *bag* in the MIML model and has m_i *instances*, i.e., $\mathbf{e_i} = \{e_i^1, e_i^2, ..., e_i^{m_i}\}$. A set of penalty labels $\mathbf{p_i}$ is associated with the bag $\mathbf{p_i}$, where each $p_i^j \in \mathbf{p_i}$ is one possible penalty value led by these events $\mathbf{e_i}$. In particular, each battery group has 24 cells with different voltage aging trend, where $\mathbf{p_i} = \{p_i^1, p_i^2, ..., p_i^{24}\}$ is a subset of all possible labels and $\mathbf{p_i} \in \mathcal{P}$. Then the objective is to learn a function $f_{MIML} : 2^{\mathcal{E}} \to 2^{\mathcal{P}}$ from a given data set $\{(\mathbf{e_1}, \mathbf{p_1}), (\mathbf{e_2}, \mathbf{p_2}), ..., (\mathbf{e_N}, \mathbf{p_N})\}$ to accurately predict the penalty label $\mathbf{p_i}$ for the bag $\mathbf{e_i}$.

Instead of receiving a set of independent labeled instances as in standard classification, MIML model receives a set of bags $\{\mathbf{e_1}, \mathbf{e_2}, ...\}$ which are labeled with the penalty values $\{\mathbf{p_1}, \mathbf{p_2}, ...\}$. Given bags obtained from different batteries at different dates, the goal is to build a classifier that will label other bags correctly. Based on the MIML algorithm, we propose a modified approach with an effective approximation of the original MIML problem. Specifically, to utilize the relations among multiple labels, our approach first learn a shared space for all the labels from the original features, and then trains label specific linear models from the shared space. To make the learning efficient, stochastic gradient descent is used to optimize an approximated ranking loss. In testing phase, our approach returns a subset of all possible labels with the prediction value, and we obtain the penalty label for each bag by selecting the one with the largest prediction value.

With the domain knowledge and our observations that the voltage is the criterion of the battery condition, the remaining lifetime can be easily predict with the voltage threshold T for pre-defined battery failure. We say the working condition of battery i is good, when the float voltage value $\mathbf{v_i}$ is higher than pre-defined threshold T. With the aging trend (v_{ia}^k, s_i^k) and penalty p^k of the voltage in time segment k, the battery remaining lifetime l_i^k for the battery i at time segment k can be calculated as $l_i^k = \frac{v_{ia}^k - T}{s_i^k + p_i^k}$.

4. EXPERIMENTS

In this section, we evaluate the prediction accuracy of the event-driven battery profiling approach using real-world trace collected from China Mobile Ltd Co and demonstrate the performance in estimating the remaining battery lifetime. To evaluate the prediction quality, we run experiment on the real-world data with 2673 batteries data, 2082 examples in the training set and 593 examples in the test set.

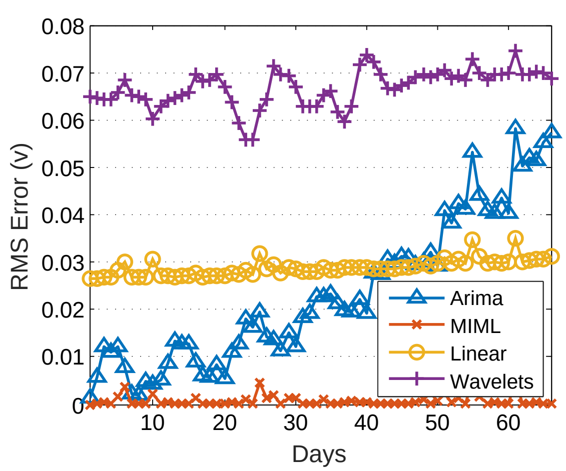


Figure 9: Root-mean-square Errors

Then we compute the root-mean-square error between the predictive and actual trends. We compare our method with ARIMA [5], Linear Regression [6] and wavelets [7], as Fig 9 shows. We can see that our method performs best among the majority of the compared schemes with the smallest root-mean-square errors, where our approach can precisely predict the battery voltage with the RMS error less than 0.01 v. Moreover, the aging trend extracted by our scheme is monotonic and satisfies the nature of the aging behavior, while LR, ARIMA, or Wavelets does not have such advantage as they have no monotonic constraint in extracting the voltage aging trend.

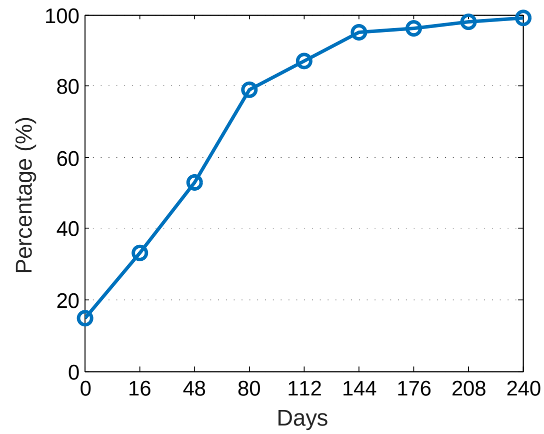


Figure 10: Survival days after our failure alert

As the battery replacement is the consequence of a battery failure [4], we choose 112 batteries with the replacement records in the log to verify the prediction accuracy for remaining lifetime. We count the survival days for each battery after the failure alert. A plot of the relation between the percentage and the survival days after our failure alert is shown in Fig. 10. We can see that 87% batteries are replaced in three months after the alters, which demonstrates our approach is a strong predictor for battery working conditions. We also observe that there are still a small number of batteries not replaced after our failure alerts. We have a close look at those batteries, and find out there are redundant battery groups in their cellular network base stations. Therefore repairing engineers postponed the maintenance service for those batteries, which also demonstrates our battery profiling approach can help the maintenance engineers to detect the potential issues in the battery groups.

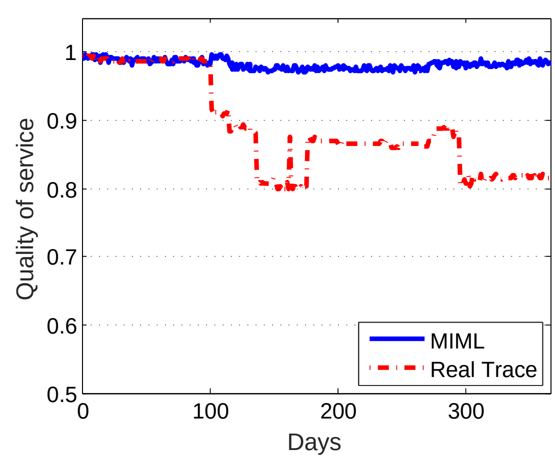


Figure 11: Service Availability

Fig. 11 plots the service availability of the base stations in the case of the power outage. We assume that the battery will be replaced after the failure alert and the voltage goes back to the normal level. The result demonstrates that the event-driven battery profiling approach together with our proposed battery maintenance and replacement scheme, can boost the cellular network service availability with an improvement of up to 18.09%.

5. RELATED WORK

The time series research has attracted significant efforts in recent decades. In this section, we highlight some relevant techniques. The most intuitive approach is smoothing, filtering and prediction, which can be done using different techniques, including ARIMA [5], Linear Regression [6], wavelets [7] and SSA [8]. Luo et al. [9] presented an approach to find the correlation between actual events and time series in order to diagnose incidents. Their approach matched the events with certain subsequences of time series in order to give a real explanation of the time series shape. Xu et al. [10] [11] combined textual messages in console logs to construct performance features and conducted the Principal Component Analysis (PCA) [12] to detect anomalies. Makanju et al. [13] proposed IPLoM by creating event descriptions based on clustering text messages in the logs, but interpreting them requires unavailable domain knowledge. Predicting failures on a data stream processing system [14] could be done by observing system's status, which uses an ensemble of decision tree classifiers and is applicable in an online setting. Finite state automata [15] can be used to model sequential dependencies between messages and detect anomalies. Most recently, Sipos et al. [4] considered log-based predictive analysis in order to monitor the conditions of the operating equipment, yet they ignored the possibility of fine temporal patterns to simplify the model. Different with the prior approaches, our event-driven battery profiling approach focuses on a large amount of battery and utilizes not only the signal value but also the events data to comprehensively predict the battery working conditions.

6. CONCLUSION

In this paper, we propose an event-driven battery profiling approach to precisely extract the features that cause the working condition degradation of the battery group. We formulate the prediction models for both battery voltage and lifetime and propose a series of solutions to yield accurate outputs, which further enable us to propose a scheme to timely schedule battery maintenance and replacement to minimize the service interruptions caused by power outages. By real world trace-driven evaluations, we demonstrate that our approach can achieve much higher prediction accuracy on the battery voltage and lifetime, which together with our proposed battery maintenance and replacement scheme, can boost the cellular network service availability with an improvement of up to 18.09%.

7. ACKNOWLEDGMENTS

This work is supported by a Canada NSERC Discovery Grant, an NSERC Strategic Project Grant, and an NSERC E.W.R. Steacie Memorial Fellowship.

Feng Wang's work is supported by a Start-up Grant from the University of Mississippi.

8. REFERENCES

[1] W. Balshe, "Power system considerations for cell tower applications," *Cummins Power Generation*, 2011.

[2] J. Vetter, P. Novák, M. Wagner, C. Veit, K.-C. Möller, J. Besenhard, M. Winter, M. Wohlfahrt-Mehrens, C. Vogler, and A. Hammouche, "Ageing mechanisms in lithium-ion batteries," *Journal of power sources*, vol. 147, no. 1, pp. 269–281, 2005.

[3] Z.-H. Zhou, M.-L. Zhang, S.-J. Huang, and Y.-F. Li, "Multi-instance multi-label learning," *arXiv preprint arXiv:0808.3231*, 2008.

[4] R. Sipos, D. Fradkin, F. Moerchen, and Z. Wang, "Log-based predictive maintenance," in *Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1867–1876, ACM, 2014.

[5] G. E. Box, G. M. Jenkins, and G. C. Reinsel, "Linear nonstationary models," *Time Series Analysis, Fourth Edition*, pp. 93–136, 1976.

[6] A. Charnes, E. Frome, and P.-L. Yu, "The equivalence of generalized least squares and maximum likelihood estimates in the exponential family," *Journal of the American Statistical Association*, vol. 71, no. 353, pp. 169–171, 1976.

[7] A. Grossmann and J. Morlet, "Decomposition of hardy functions into square integrable wavelets of constant shape," *SIAM journal on mathematical analysis*, vol. 15, no. 4, pp. 723–736, 1984.

[8] J. B. Elsner and A. A. Tsonis, *Singular spectrum analysis: a new tool in time series analysis*. Springer Science & Business Media, 2013.

[9] C. Luo, J.-G. Lou, Q. Lin, Q. Fu, R. Ding, D. Zhang, and Z. Wang, "Correlating events with time series for incident diagnosis," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1583–1592, ACM, 2014.

[10] W. Xu, L. Huang, A. Fox, D. A. Patterson, and M. I. Jordan, "Mining console logs for large-scale system problem detection.," *SysML*, vol. 8, pp. 4–4, 2008.

[11] W. Xu, L. Huang, A. Fox, D. Patterson, and M. Jordan, "Online system problem detection by mining patterns of console logs," in *Data Mining, 2009. ICDM'09. Ninth IEEE International Conference on*, pp. 588–597, IEEE, 2009.

[12] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2002.

[13] A. A. Makanju, A. N. Zincir-Heywood, and E. E. Milios, "Clustering event logs using iterative partitioning," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1255–1264, ACM, 2009.

[14] X. Gu, S. Papadimitriou, P. S. Yu, and S.-P. Chang, "Online failure forecast for fault-tolerant data stream processing," in *Data Engineering, 2008. ICDE 2008. IEEE 24th International Conference on*, pp. 1388–1390, IEEE, 2008.

[15] Q. Fu, J.-G. Lou, Y. Wang, and J. Li, "Execution anomaly detection in distributed systems through unstructured log analysis," in *Data Mining, 2009. ICDM'09. Ninth IEEE International Conference on*, pp. 149–158, IEEE, 2009.