SHARING POWER STORAGE FOR UNDEVELOPED COUNTRIES

ZISHUO ZHAO, 2015011255

Abstract. Motivated to provide stable power to undeveloped countries, in which the grid can only keep on for a short time each day, an organization can be set up to provide pooled power storage for users. To make the system efficient, we proposed a mechanism to decide on the capacity and charge users for their share of the total cost. Differentiated from usual ideas to maximize profit, this paper is designed for non-profit organizations, focusing on maximizing the social welfare while charging users in a fair way to cover the cost. The decision of the capacity is based on sampling—the organization randomly power on part of all users for demand data with which the capacity is decided on, and the optimal size of the sample is also roughly estimated with statistics to trade off between sampling cost and deviation from optimal capacity due to sampling errors. Afterwards, the charging plan is designed for fairness, and in distinct perspectives, two slightly different policies are worked out, both charging for the total consumption and demand variance, and covering the total cost. When the total number of users is large, it is also proved that the both policies converge to a simplest idea to charge users proportional to their total consumption of power, which is carried out in most modern places.

1. Introduction

In the modern world, electricity is essential for daily life of citizens. However, in some undeveloped countries, it is even beyond imagination to have power on for 24 hours a day—although one hour or two is possible. For convenience, they may buy or rent batteries to store power and have lights on when power is off.

However, it is hard for a family to decide how much power is needed each day, and generally, the demand varies in different time. If every house buys or rents a battery for maximum possible demand, at most time some capacity is spared, leading to insufficient utilization of resource—especially for poor families even strugging for the device. Therefore, by sharing the storage, the risk can be pooled and the capacity can be more efficiently utilized.

In this paper, we designed a scheme to decide on the optimal capacity and a corresponding policy for the storage sharing project. In the scheme, the organization firstly randomly powers on $O(n^{\frac{2}{3}})$ users to get informations about the energy comsumption among n users. Based on the information, the organization decides on the capacity and the detailed policy of charging.

2. System model

2.1. **Description.** In an undeveloped city, the power is on only for a few hours a day. To satisfie users' demands for full-day electricity supply, an organization intends to set up a storage device with capacity C and daily cost s. When the power is on, the organization charges the storage to full, and when the power is off, it provides power to users with the capacity and charge them some money.

Here, we have some assumption:

(1) The demand of user i is an i.i.d random variable $X_i \sim N(\mu_i, \sigma_i^2)$, in which μ_i and σ_i are drawn from some uniform distributions.
(2) All demands do not vary with seasons. (This may be oversimplified, but in practice, the organization may undertake different policies according to seasons.)
(3) If power shortage happens, the storage is allocated proportional to demands, but there is an extra cost proportional to exceeding underfall, with cost rate $\gamma \gg s$.
(4) Because shortage probability is low, when computing users' consumption we ignore the shortage case.

2.2. **Calculation for optimized sampling policy.** If there are n users, then as their consumption are independent, and we know all information of their consumption distribution, we get

$$\sum_{i=1}^{n} X_i \sim N(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2).$$

for each day, we denote the total demand $S = \sum_{i\in[n]} X_i$, then the opportunity cost is

$$c_{op}(C_{emp}) = \begin{cases} s(C_{emp} - S), & S \leq C_{emp} \\ \gamma(S - C_{emp}), & S > C_{emp} \end{cases}.$$

To make the shortage probability below ϵ, we only need to make the capacity

$$C \geq \sum \mu_i + K_1 \sqrt{\sum_{i=1}^{n} \sigma_i^2},$$

in which the tail probability

$$Q(K_1) = \frac{s}{\gamma + s}$$

based on the result of news-vendor model.

The easiest way to decide on the optimal C is to gather the μ_i and σ_i from each user. But in practice, users may report strategically to maximize their own utility, or more likely—to have no concept of their electricity demand, because they have never enjoyed the convenience of 24-hour power before! Therefore, we can experimentally power on part of users as an experiment, and get useful informations for final design.

2.2.1. *Asymptotic analyses.* The optimal capacity is

$$C_* = \sum_{i=1}^{n} \mu_i + K_1 \sqrt{\sum_{i=1}^{n} \sigma_i^2},$$

and we define

$$C_d = n \cdot E[\mu] + K_1 \sqrt{n \cdot E[\sigma_i^2]}.$$

By central limit theorem, for $n \to \infty$, we can approximately get

$$\sum_{i=1}^{n} \mu_i \sim N(n \cdot E[\mu], n \cdot Var[\mu]).$$

By 3σ principle, it is almost sure that

$$\left| \sum_{i=1}^{n} \mu_i - nE[\mu] \right| < 3\sqrt{n \cdot Var[\mu]},$$

and similarly, it is almost sure that

$$\left| \sum_{i=1}^{n} \sigma_i^2 - nE[\sigma^2] \right| < 3\sqrt{n \cdot Var[\sigma^2]}.$$

Therefore, we can approximate $\frac{1}{n}\sum_{i=1}^{n} \mu_i$ and $\frac{1}{n}\sum_{i=1}^{n} \sigma_i^2$ with $E[\mu]$ and $E[\sigma_i^2]$ respectively, with roughly $\frac{1}{\sqrt{n}}$ error.

Now we consider a fixed random subset $M \in [n]$ with $|M| = m$, for $m \to \infty$ and $m = o(n)$, by central limit theorem,

$$E_{i\in M}[\mu_i] \sim N(E[\mu], \frac{Var[\mu]}{m})$$

$$E_{i\in M}[\sigma_i^2] \sim N(E[\sigma_i^2], \frac{Var[\sigma_i^2]}{m}).$$

Based on $E_{i\in[M]}[\mu_i]$ and $E_{i\in[M]}[\sigma_i^2]$, the organization decides on the capacity

$$C_{emp} = nE_{i\in M}[\mu_i] + K_2\sqrt{nE_{i\in M}[\sigma_i^2]}.$$

So

$$Var[nE_{i\in M}[\mu_i]] = n^2\frac{Var[\mu]}{m}$$

$$Var\left[K_1\sqrt{nE_{i\in M}[\sigma_i^2]}\right] \approx \frac{n^2 Var[\sigma_i^2]}{4m \cdot E^2[\sqrt{n\sum_{i\in m}\sigma_i^2}]}.$$

For $n \to +\infty$, the influence of fluctuation of $K_1\sqrt{nE_{i\in M}[\sigma_i^2]}$ is negligible, and $\left|\sqrt{Var[A+B]} - \sqrt{Var[A]}\right| \leq \sqrt{Var[B]}$ so we have

$$Var[C_{emp}] = n^2\frac{Var[\mu]}{m}(1+o(1))$$

$$\begin{aligned} E[(C_{emp} - C_*)^2] &= E[(C_{emp} - C_d + O(\frac{1}{\sqrt{n}}))^2] \\ &= Var[C_{emp}](1+o(1)) \\ &= n^2\frac{Var[\mu]}{m}(1+o(1)). \end{aligned}$$

2.2.2. *Arguments for sampling.* In fact, for $m \ll n$, with high probability the difference between C_* and C_d is negligible compared to the fluctuation of C_{emp}, so when considering the cost due to the error of C_{emp}, we can assume $C_* = C_d$.

To get correct usage data for the sample, we should provide them with enough power—more than the optimal capacity. For early construction and data analysis for the sample, there will be some cost, roughly proportional to the sample size m. On the other hand, the larger sample will obtain a more accurate estimation of parameters and less cost of policy deviation, and it has been proved that the range of sampling error is roughly $\frac{n}{\sqrt{m}}$, while the deviation from users' random consumption is proportional to $\sqrt{n}$.

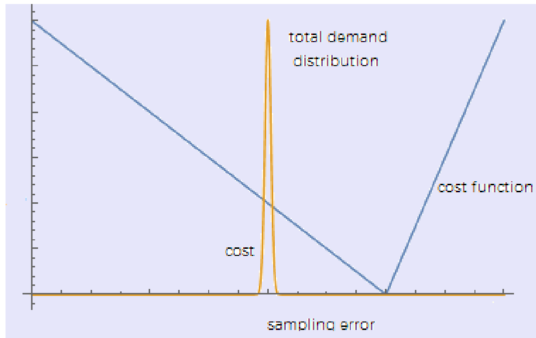
Firstly, we consider a case where $\sigma_i = 0$ for any i. In this case, the total demand is fixed, so the error cost is proportional to the error of C_{emp}. Therefore, the total cost of sampling process and sampling error is

$$c_{spl} = \Theta(m) + \Theta(\frac{n}{\sqrt{m}}),$$

we can get

$$m_{opt} = \Theta(n^{\frac{2}{3}}).$$

Besides, in general cases, for $m = o(n)$ and $n \to \infty$, $\frac{n}{\sqrt{m}} \gg \sqrt{n}$, so we can still regard the total demand as fixed (as in the figure below). Therefore, generally, the optimal sample size is $\Theta(n^{\frac{2}{3}})$.



In practice, the exact size of optimal sampling varies with many aspects. For example, it decreases when sampling cost increases, and increases when consumption varies among users more widely.

2.2.3. *Special sampling case.* In general cases where $n \to +\infty$, the optimal m increases in the order of $n^{\frac{2}{3}}$. This result is based on the fact that when $n \to +\infty$ and $m = o(n)$, the sampling error takes the main part. However, when n is not so large, we cannot ignore the effect of demand fluctuation, the result can be different.

Now we consider another extreme case, where the demand fluctuation takes the main part. Then, the sampling error leads to a small deviation from the optimal fractile in the SPP model. Take $\frac{C_{emp}-C_*}{\sqrt{\sum_{i=1}^{n}\sigma_i^2}} \to 0$, the CDF can be approximated to be linear around C_*, then the cost of sampling error is in the order of $(\frac{n}{\sqrt{m}})^2 = \frac{n^2}{m}$. In this case, we have

$$m_{opt} \sim n.$$

Here we have characterized the pattern of optimal sample size m. For fixed distributions, when n is small, m is roughly proportional to n, and when n is large enough, m increases with $n^{\frac{2}{3}}$.

3. Policy for payment

In previous parts we have only discuss about how to decide on the (approximately) optimal storage capacity. In practice, it costs much money for setting up the device. In this part, we discuss about how much money should be charged from each user i.

Here we have some assumptions:

(1) The organization knows the distribution of μ_i and σ_i accurately, and the capacity decided on (C_{emp}) is the optimal value C_*, ignoring errors discussed in previous parts. (In fact, C_{emp} is a first-order unbiased estimation of C)
(2) $E_{i\in[n]}[\mu_i] = E[\mu]$, $E_{i\in[n]}[\sigma_i^2] = E[\sigma_i^2]$, implying $C_* = C_d$.
(3) All costs are amortized to every day, and daily cost is $s \cdot C_*$, proportional to the capacity.
(4) We only discuss about the cost of device. The cost of electricity itself is evidently paid by each user according to consumption.
(5) Users' demands are not affected by prices, and the organization does not mean to make profit. It charges only to cover the cost.

The general idea is to charge users for the part of devices that they demand, although the actual policy can be a little different. because adding new users leads to better risk pooling and better utilization, while each user should be charged in a fair way, while the organization should have its costs covered.

3.1. **The payment policy based on contribution.** In the policy, users are charged for every $T-$day period. We assume that in each period all μ_i and σ_i remains constant, while in different periods they can vary under the restriction Assumption 2. During the $T-$day period, we compute all μ_i and σ_i^2 with formulae

$$\hat{\mu_i} = \frac{1}{T}\sum_{t=1}^{T} X_i(t)$$

$$\hat{\sigma_i^2} = \frac{1}{T-1}\sum_{t=1}^{T}(X_i(t) - \mu_i).$$

Here, $\hat{\mu_i}$ and $\hat{\sigma_i^2}$ are unbiased estimations of μ_i and σ_i^2. In following calculations, we regard them as the accurate value, and the results have no first-order bias.

Amortized in each day, the cost of the device is

$$\begin{aligned} W &= s \cdot C_* \\ &= s \cdot \left(\sum_{i=1}^{n} \mu_i + K_1 \sqrt{\sum_{i=1}^{n} \sigma_i^2} \right). \end{aligned}$$

Considering the contribution of user i to W,

$$\frac{\partial W}{\partial \mu_i} = s$$

$$\frac{\partial W}{\partial \sigma_i} = s \cdot K_1 \frac{\sigma_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}.$$

A first order approximation of user i to W is

$$w_i = s \cdot \left(\mu_i + K_1 \frac{\sigma_i^2}{2\sqrt{\sum_{i=1}^{n} \sigma_i^2}} \right).$$

However, when we add all w_i up,

$$\begin{aligned} \sum_{i=1}^{n} w_i &= s \cdot \sum_{i=1}^{n} \left(\mu_i + K_1 \frac{\sigma_i^2}{2\sqrt{\sum_{j=1}^{n} \sigma_j^2}} \right) \\ &= s \cdot \left(\sum_{i=1}^{n} \mu_i + \frac{K_1}{2} \sqrt{\sum_{i=1}^{n} \sigma_i^2} \right) \\ &< W. \end{aligned}$$

This is due to the risk-pooling effect. When we share more, the capacity can be more efficiently utilized, and the marginal cost for new users decreases. But the organization should have its costs covered, so we have the daily payment adjusted as following:

$$w_i^{(1)} = s \cdot \left(\mu_i + K_1 \frac{\sigma_i^2}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}} \right)$$

then the cost-covering condition can be satisfied:

$$\sum_{i=1}^{n} w_i^{(1)} = W.$$

Although slightly altered, this payment rule still reflects the relationship between usage and cost: the contribution of one user to the total cost consists of two parts— the first part is the basic storage, which equals to its expected demand; the second part is risk-reserving storage, which is proportional to the square of the standard deviation of its demand.

3.2. **An alternative policy considering Opportunity Fairness.** In the policy above, User 1 may argue: I use exactly twice as much electricity as User 2, but my variance cost is 4 times of x. That is unfair!

In the organization's perspective, the policy is fair. Although User 1 uses twice as much electricity as User 2, the organization has to set up more than twice the capacity for User 1 than User 2, so charging more is fair. However, in users' perspective, under the Veil of Ignorance, the behavior of User 1 is indifferentiable from two identical User 2. The difference of payment is only due to the syncrony of the two, which brings more cost to the organizer because it is bad for risk-pooling.

Has User 1 done anything wrong? No. The extra cost is due to the sharing mechanism, not the user. So in the perspective of users, or the fairness of outcome, User 1 should pay twice as much as User 2. Nevertheless, it is still true that more deviation leads to more cost, so the cost for risk-reserving storage should be paid by users, but in this perspective, it is more fair to be allocated proportional to σ_i, not σ_i^2.

Therefore, user i should pay

$$w_i^{(2)} = s \cdot \left(\mu_i + K_1 \frac{\sigma_i \sqrt{\sum_{i=1}^n \sigma_i^2}}{\sum_{i=1}^n \sigma_i} \right)$$

for each day.

3.3. **Conclusion.** In this part, we have derived three payment policies. w_i is the shadow price of the entrance of i, but due to the scale effect of the system, after removing some users, the shadow prices of other users increase, so this policy cannot cover the total cost; $w_i^{(1)}$ is based on the shadow price and observes the fairness of outcome, and $w_i^{(2)}$ still considers the cost of expected demand and variance but respects the fairness of opportunity more. All policies are fairly reasonable in some way, but none of them are "perfect". Generally speaking, the conflict of the "fairness of outcome" and "fairness of opportunity" is an eternal conflict in society.

However, we can notice that when $n \to \infty$, the cost of risk pooling for every person is diminishing, and w_i, $w_i^{(1)}$, $w_i^{(2)}$ all converge to $s \cdot \mu_i$. It shows that when we share more, there will be not only more efficiency, but also fewer conflicts.

4. Acknowledgements

This project is mainly based on lecture materials and my knowledge in statistics, so I have not directly referred to literature. Nevertheless, in the process of completing this project, I did gain a lot of inspiration and wisdom from teachers (Chenye Wu and Yang Yu), and many classmates.