Solar Radiation Prediction and Energy Allocation for Energy Harvesting Base Stations

Yanan Bao∗, Xiaolei Wang†, Xin Liu∗, Sheng Zhou† and Zhisheng Niu†
∗Dept. of Computer Science, University of California, Davis 95616
Email: ynbao@ucdavis.edu, liu@cs.ucdavis.edu
†Dept. of Electronic Engineering, Tsinghua University, Beijing, 100084
Email: wang-xl09@mails.tsinghua.edu.cn, sheng.zhou@tsinghua.edu.cn, niuzhs@tsinghua.edu.cn

Abstract—In this paper, we study how to use the solar radiation model to predict energy arrivals and to allocate energy resource at an energy harvesting base station (BS). First, some primary knowledge about solar radiation is reviewed and summarized. We present two solar energy models for cloudless days and cloudy days, respectively. Then artificial neural network (ANN) is used to predict solar energy arrivals in a short period, which has an improved performance compared with the previous linear model. In the end, the allocation of received energy is considered, and one optimal offline algorithm and four heuristics online algorithms are proposed. We evaluate the performance of the algorithms using Denver’s solar radiation data in recent 27 years from National Renewable Energy Laboratory (NERL). Simulation results show our prediction and optimization algorithm achieves nearly optimal performance.

I. INTRODUCTION

According to a survey conducted by International Telecommunication Union (ITU) in 2003, more than 2.5 billion people (which is 40% of the total population) still live in remote and rural areas with limited communication access [1]. Worse more, electricity supply in most rural areas is always insufficient or even non-existent. Therefore, how to power the BSs in rural areas is a challenging issue facing developing countries and even the world.

To address this challenge, energy harvesting has been introduced into communication industry as the power source of BSs in rural areas. Now BSs that harvest renewable energy, such as solar and wind, are gradually deployed in some rural areas throughout the world. For instance, China Mobile had established the world’s largest solar energy BS cluster in Tibet [2]. Moreover, among the 1000+ BSs in Tibet, about 80% of them are powered by solar energy. This reveals the potential of energy harvesting techniques to be applied in rural communications and future’s green communications.

If the traffic has no large variations, energy output of BSs should be stable to guarantee high QoS and high energy efficiency. In this work, we consider that BSs only use the harvested energy and have no alternative energy source to back up. In this case, the prediction of future energy arrivals is important, since it can balance the energy storage and energy output in the temporal domain. More accurately, traffic modeling will be considered in our future work, such that energy output matches the traffic. Notably, when the battery capacity is large enough, we can use a constant energy output rate that equals to the average energy input rate. However, batteries with very large capacity increase the deployment cost dramatically. Therefore, with limited battery capacity, good energy allocation algorithms are needed to best allocate the harvested energy.

In this paper, we only consider BSs equipped with solar panel. Solar energy models have been systematically studied by geographers and meteorologists since 1970s. However, most models are too complex to be applied, and not all of the parameters are always available. For instance, in the model contributed by [3]–[6], air pressure, temperature, water vapor, oxygen, ozone, aerosol and spectrum are all included. Therefore, our work begins by presenting a simple and general solar energy model.

For resource scheduling in energy harvesting communications, much work has been done either in different scenarios, or with different targets, or in different methods. For instance, [7] proposes energy-neural mode that can use as much energy as possible from the environment in sensor networks. [8] presents optimal energy management schemes in energy harvesting systems with fading channels to maximize the number of bits sent within a deadline. Moreover, the duration that all the packets are delivered is minimized by changing transmission rate according to traffic load and available energy in [9], and the short-term throughput is maximized for a link with an energy harvesting transmitter in [10]. However, most of the work just assumes random energy arrivals and doesn’t take into account the energy arrival prediction, which will be solved in our work.

The main contributions of this work are:

• We propose a simple and general solar energy model that is more precise compared with the previous linear model;
• Several energy allocation algorithms are proposed and performance evaluation shows the prediction and optimization algorithm achieves nearly optimal utility;
• We extensively evaluate the performance of the proposed models and algorithms using 27-year history data [11] at Denver provided by NERL.

The remainder of this paper is organized as follows. In Section II, we introduce the solar energy model for cloudless days and use the past 27-year data in Denver to test the model. Section III presents our solar energy model for cloudy days and it is compared with the Sharma model. Section IV includes
several algorithms in the energy allocation problem with or without solar energy prediction. Performance evaluation is given in Section V. Finally, Section VI concludes the paper and discusses the future work.

II. SOLAR RADIATION MODEL FOR CLOUDLESS DAYS

Solar radiation is related to many parameters that may be location-dependent. In this section, we hope to highlight the relation between solar radiation and the key parameters, including extraterrestrial solar irradiation, cycle of year, cycle of day, latitude, longitude and altitude. Notably, these parameters are all weather-independent.

The sun is located at the foci of an elliptical revolved by the earth. The average distance from sun to earth is denoted by \( r_0 = 1.496 \times 10^8 \text{km} \) which is called one astronomical unit. The maximal energy received on the surface of the earth in the extraterrestrial space is \( 1367 \text{ w/m}^2 \) and it is denoted by \( I_0 \) [12].

Day angle \( \Gamma \) is defined as

\[
\Gamma(d) = 2\pi(d - 1)/365, \tag{1}
\]

where \( d \) is the day number in a year [12]. Here a year is assumed to always have 365 days and February has 28 days. \( E_0 \) stands for the ratio that solar radiation can reach the earth, which is determined by the distance \( r(\Gamma) \) from the earth to the sun. The benchmark is taken on the circle with radius \( r_0 \). According to the work of Spencer [13], \( E_0 \) can be expressed by a Fourier series as follows:

\[
E_0 = \frac{r_0^2}{r} = 1.000110 + 0.034221 \cos(\Gamma) + 0.001280 \sin(\Gamma) + 0.000719 \cos(2\Gamma) + 0.000077 \sin(2\Gamma). \tag{2}
\]

The declination angle \( \delta \) is the angle between the sun-earth line and the ecliptic plane on which the earth revolves around the sun. The following expression gives an estimation of \( \delta \) with a maximum error of 0.0006 rad [13].

\[
\delta = 0.0006918 - 0.399912 \cos(\Gamma) + 0.070257 \sin(\Gamma) - 0.006758 \cos(2\Gamma) + 0.000907 \sin(2\Gamma) - 0.002697 \cos(3\Gamma) + 0.00148 \sin(3\Gamma). \tag{3}
\]

The length of each day is not exactly 24 hours, which means the local meridian is slightly changing everyday. This causes a time discrepancy that can goes as much as 16 minutes [12]. Also according to [13], the discrepancy can be expressed as follows:

\[
E_t = \frac{229.18}{60} \left( 0.000075 + 0.001868 \cos(\Gamma) - 0.032077 \sin(\Gamma) - 0.014615 \cos(2\Gamma) - 0.04089 \sin(2\Gamma) \right). \tag{4}
\]

Solar radiation may be recorded in terms of local apparent time (LAT) or local standard time (LST). We use LAT in the rest of this paper. The relation between LAT and LST is given as:

\[
\text{LAT} = \text{LST} + \frac{4(L_s - L_e)}{60} + E_t, \tag{5}
\]

where \( L_s \) and \( L_e \) are the longitude of the observer and the longitude of the time zone, respectively.

The zenith angle \( \theta_z \) is the angle between sun-observer line and the horizontal ground, and at hour \( t \) it is as

\[
\cos \theta_z(t) = \sin \delta \sin \omega + \cos \delta \cos \omega \cos(\pi(t - 12)/12), \tag{6}
\]

where \( \omega \) is the latitude of the observer. Hottel Model [14] gives the total solar radiation at time \( t \) on a horizontal plane as

\[
E_{\text{max}}(t) = \left\{ \begin{array}{ll}
I_0E_0(0.2710 + 0.7061(a_0 + a_1 \cos(\frac{t}{12}))) & \text{if } \cos \theta_z > 0; \\
0 & \text{otherwise.} 
\end{array} \right. \tag{7}
\]

\( E_{\text{max}}(t) \) includes both direct solar radiation and diffuse irradiance. Constants \( a_0, a_1 \) and \( k \) are related to the altitude \( h \) in kilometers. In different regions, \( a_0, a_1 \) and \( k \) may be slightly different, but general settings are enough in this work. Here we have

\[
a_0 = 0.4237 - 0.00821(6 - h)^2; \tag{8}
\]
\[
a_1 = 0.5055 + 0.00595(6.5 - h)^2; \tag{9}
\]
\[
k = 0.2711 + 0.01858(2.5 - h)^2. \tag{10}
\]

To test the performance of this model, we use the solar radiation data from the Measurement and Instrumentation Data Center (MIDC) of NERL. The database provides irradiance and meteorological data from about 30 stations around America for decades. We choose 27 years’ data from Denver as an example. The data is from the Measurements and Instrumentation Data Center (MIDC), National Renewable Energy Laboratory (NREL) (http://www.nrel.gov/midc/). Note that the solar radiation we calculate is the maximal energy that can be received at any time of a day. By integrating on variable \( t \) in Eq. (7), we can get the hourly incoming energy.

We compare hourly received energy generated by this model and the processed historical data in Fig. 1. The different lines correspond to different hours during daytime. Several steps were done before we get \( E_p \), such that it is smoothed and has less vibration. Denote the energy arrival at hour \( h \), day \( d \) and year \( y \) by \( E_{h,d,y} \).

\[
S(h,d) = \{E_{h,d+i,j} : i = -m, -m + 1, \cdots , m, \quad j = 1,2,\cdots,27 \}. \tag{11}
\]

Choose \( E_{p}(h,d) \) as the \( l \)-th largest number in set \( S(h,d) \). Here we choose \( m \) to be 5 and \( l \) to be 11. When \( m \) and \( l \) are
small, random noise distorts the lines. However, when \( m \) and \( l \) get much larger, distortion occurs due to over smoothing. Through this process, complicated weather conditions can be screened and the data can be properly smoothed. Fig. 1 shows this model can match the real solar radiation in clear sky days very well.

III. SOLAR RADIATION MODEL FOR CLOUDY DAYS

Research shows a strong negative correlation between received solar energy and the concurrent cloud coverage. We assume the short-term cloud coverage is available from weather forecast, which can be used to predict future energy arrivals. In [15] Sharma et al. proposes that

\[
E_c(t) = E_{\text{max}}(t)(1 - p_c(t)), \tag{12}
\]

where \( p_c(t) \) is the percentage of cloud coverage at time \( t \) and \( E_c(t) \) is the solar radiation considering cloud. We obtain two cloud parameters from MIDC: one is the opaque cloud coverage and the other is the total cloud coverage. Linear regressions on these two parameters show using opaque cloud coverage has less root mean square error (RMSE) and thus it is used to make prediction in the rest of this paper.

Define energy reception ratio at time \( t \) by \( r(t) = E(t)/E_{\text{max}}(t) \), where \( E_{\text{max}}(t) \) is defined in Eq. (7). Linear regression based on the data from Denver in 2010 shows \( r = 0.9987 - 0.9417p_c \). This is very close to \( r = 1 - p_c \), which is proposed by [15]. We call this model simple linear model.

The histogram of daytime \( r(t) \) and the autocorrelation of \( r(t) \) are shown in Fig. 3. The large spike at \( r(t) = 0 \) in Fig. 3(a) is mainly caused by data loss and device error. The main part ranges from 0.1 to 1.1: it is evenly distributed from 0.1 to 0.9 and is approximately Gaussian distributed from 0.9 to 1.1. Fig. 3(b) shows data within 12 hours and from a 24-hour cycle is highly correlated. The reason is that even though the main periodicity has been taken into account by \( E_{\text{max}} \), the 24-hour cycle still has effects on \( r(t) \) by other weather parameters, such as temperature and pressure.

We enhance the model in Eq. (12) as

\[
E(t) = E_{\text{max}}(t)(1 - p_c(t))(1 + v(t)). \tag{13}
\]

Here \( v(t) \) is a variable with no relation to the opaque cloud coverage. Statistics show the histogram of \( v(t) \) and the autocorrelation of \( v(t) \) in Fig. 4. Fig. 4(a) shows \( v(t) \) resembles Laplace distribution with zero mean. Fig. 4(b) shows \( v(t) \) has a strong correlation within only 1 hour, which means \( v(t) \) is much more independent than \( r(t) \) in the temporal domain.

Artificial neural network (ANN) is used to predict \( v(t) \) and further to predict the energy arrival \( E(t) \), which is compared with the simple linear model in Eq. (12). The linear model only use the current cloud coverage, while ANN model uses the past
v(t) as inputs. There are 10 hidden layers and 12 feedbacks in the ANNs. The data from Denver goes from 2008 to 2012. Linear model has a RMSE of 66.7236 W/m² in the last 2 years. For the ANN model, the first 3-year data is taken to train the network by Bayesian regularization backpropagation, and the last 2-year data is taken to test the network. It is shown in Fig. 5 that in 12 hours, ANN has less RMSE, but for a prediction longer than 12 hours, ANN is actually worse than the simple linear model. For instance, RMSE of the prediction using ANN is reduced by nearly 10% for the next 1 hour. This also verifies that v(t) is less autocorrelated in the temporal domain and linear model is actually an unbiased prediction. Therefore, in our prediction, ANN is used within 12 hours and linear model is used beyond 12 hours. We call this model ANN-linear hybrid model.

IV. ENERGY ALLOCATION ALGORITHMS

We consider a BS using solar panel to harvest solar energy and having a battery to store the solar energy. With different kinds of solar panel, the amount of energy that can be harvested is slightly different. Here we assume the solar radiation can be fully harvested by the solar panel and the transfer efficiency to be 1 for notation simplicity. The battery has its capacity of $E_0$. In general, $E_T$ equals to the total arrived energy in 1 to 3 days. The BS needs to allocate the stored energy into the next slots. Here one slot corresponds to 1 hour.

Since BSs require stable energy output to guarantee QoS, our utility function is defined as a concave increasing function:

$$G(u_i) = \log(e + u_i/u_0),$$

where $u_i$ is the energy output at hour $i$, $e$ is a small constant and $u_0$ is the benchmark for energy output. $u_0$ is the typical energy output of a BS and when $u_i = (1-e)u_0$, the utility $G(u_i) = 0$. Here we assume the traffic is greedy and channel fading is time invariant, such that the utility is only determined by the energy output $u_i$. This function on variable $u_i$ can give large penalty as $\log(e)$ when $u_i$ approaches $0$ and logarithm utility when $u_i > u_0$. Our objective is to maximize the average utility for a long time, i.e.,

$$\lim_{T\to\infty} \frac{\sum_{j=1}^{T} \log(e + u_i/u_0)}{T}.$$  

Next we introduce one offline algorithm and four online algorithms to allocate the harvested energy. The offline performance gives the upper bound for the online algorithms.

A. Offline algorithm

Assume the future solar energy arrivals are known in advance. Denote the energy arrivals at hour $1, 2, \ldots, T$ by $\vec{e} = [e_1, e_2, \ldots, e_T]$. $\vec{u}$ is a vector contains the energy output $u_1, u_2, \ldots, u_T$. The optimization function is

$$\max \sum_{t=1}^{T} \frac{(\log(e + u_t/u_0))}{T};$$

s.t. $E_0 + C_a\vec{e} - C_b\vec{u} \geq 0$;

$E_0 + C_b(\vec{e} - \vec{u}) \leq E_1$;

$\vec{u} \geq 0$.

Here both $C_b$ and $C_a$ are $T \times T$ matrixes of 0s and 1s:

$$C_b(i, j) = \begin{cases} 1, & \text{if } i \leq j; \\ 0, & \text{otherwise}. \end{cases}$$  

and

$$C_a(i, j) = \begin{cases} 1, & \text{if } i < j; \\ 0, & \text{otherwise}. \end{cases}$$

$C_b$ and $C_a$ show our assumption that only the previously harvested energy can be used in this hour and the energy harvested in this hour will be stored for next hours. Since this problem has a concave objective function and linear constraints, it can be solved by convex optimization.

However, in reality, even though weather forecast provides sufficient cloud coverage information in one or two days. Randomness still exists and perfect energy arrival prediction is difficult. This is due to two reasons: one is the existence of other parameters that affect solar harvesting in short term; the other is the imprecision of weather forecast in long term.

B. The long term average

If the battery has infinite capacity, the BS should set a constant energy output rate that equals to the average energy input rate. The utility function can be maximized with this simple setting. This is the intuition behind the long term average algorithm, in which

$$u_i = \frac{\sum_{j=1}^{T} e_j}{T}, \quad i = 1, 2, \ldots, T.$$  

on the other hand, when battery capacity is limited, this algorithm may run out of energy, resulting in lower utility.

C. Conservative algorithm

In this algorithm, running out of energy is highly forbidden due to the severe penalty. This policy assumes the worst case for next $N$ hours, i.e., $\vec{e} = 0$. In order to deal with this case, the BS has to allocate its energy evenly on the next $N$ hours as shown in Algorithm 1. This conservative algorithm tends to use less energy most of the time.

It is easy to see the battery keeps in a high level in this algorithm. This may limit the potential of energy harvesting.
D. Prediction algorithm

The above algorithm is too conservative in considering the future energy arrivals. An algorithm including energy arrival prediction is proposed, as depicted in Algorithm 2. We assume weather forecast can provide cloud coverage in the next $L$ hours and therefore energy arrivals in recent $L$ hours can be predicted by the ANN-linear hybrid model in Sec. III. The long term energy arrivals in $N$ hours are predicted by Eq. (7). Here we multiply $E_{\text{max}}(i)$ by the average energy reception ratio $\bar{r}(i)$.

Algorithm 2 Prediction algorithm

Input:
Initial battery level $E_0$, battery capacity $E_l$, average energy reception ratio $\bar{r}(i)$, energy allocation length $N$, cloud forecast length $L$, time length $T$;

Output:
1: $i = 0$;
2: while $i < T$ do
3: predict the energy income $e_{i+1}, e_{i+2}, \ldots, e_{i+L}$ by ANN-linear hybrid model;
4: compute the future energy income $e_j = E_{\text{max}}(j)\bar{r}(i)$, for $j = i + L + 1, i + L + 2, \ldots, N$;
5: $u(i) = (E_0 + \Sigma_{j=1}^{N} e_j)/N$;
6: record the real energy income $e_r(i)$;
7: $E_0 = \min\{E_0 - u(i) + e_r(i), E_l\}$;
8: $i = i + 1$;
9: end while

E. Prediction and optimization algorithm

At last, we propose an algorithm that first predicts solar energy arrivals and then optimizes energy allocation. Given a prediction about energy arrivals in the next $N$ hours, the energy output $\bar{u} = [u_1, u_2, \ldots, u_N]$ can be optimized in problem (16). Note that here $N$ replaces $T$ in problem (16). Here we use only $u^*_1$ as the energy output at this time and drop $u^*_2, u^*_3, \ldots, u^*_N$. Then make a new prediction for the next $N$ hours and optimize $\bar{u}$ iteratively. This algorithm is depicted in Algorithm 3.

Note that the prediction may not be precise due to all kinds of reasons. This will cause the performance degradation compared with the optimal offline algorithm.

Algorithm 3 Prediction and optimization algorithm

Input:
Initial battery level $E_0$, battery capacity $E_l$, average energy reception ratio $\bar{r}(i)$, energy allocation length $N$, cloud forecast length $L$, time length $T$;

Output:
1: $i = 0$;
2: while $i < T$ do
3: predict the energy income $e_{i+1}, e_{i+2}, \ldots, e_{i+L}$ by ANN-linear hybrid model;
4: compute the future energy income $e_j = E_{\text{max}}(j)\bar{r}(i)$, for $j = i + L + 1, i + L + 2, \ldots, N$;
5: solve the optimization problem (16) (by replacing $T$ with $N$) and get the optimal solution $\bar{u}^* = [u_1^*, u_2^*, \ldots, u_N^*]$;
6: $u(i) = u_1^*$;
7: record the real energy income $e_r(i)$;
8: $E_0 = \min\{E_0 - u(i) + e_r(i), E_l\}$;
9: $i = i + 1$;
10: end while

V. Performance Evaluation

We compare the proposed algorithms using the first 15 days’ solar radiation data in 2010. The parameter settings are as follows: $\epsilon = 0.1$, $u_0 = 100$ Wh, $\bar{r}(i) = 0.8$, $T = 24 \times 530$ h, $N = 24 \times 5$ h and $L = 24$ h. Fig. 6 and Fig. 7 show the snapshots of energy output and battery level in the 15 days. Here we assume that $E_l$ equals to the average energy arrivals in 24 hours and the initial state of the battery is full. We use the unit of hour to denote the battery capacity: 1 hour means the battery can store the average energy arrivals in 1 hour.

Our proposed prediction and optimization algorithm achieves good performance by avoiding energy overflow and energy outage. For example, at about hour 16 in Fig. 7, three online algorithms except the prediction and optimization algorithm all waste some energy due to the energy overflow. The prediction and optimization algorithm that has a spike before hour 16 just bursts forth some energy by this spike in advance. The battery level of the long term average algorithm reaches the battery capacity several times in 15 days, which means it has the worst performance to avoid energy overflow. Both the long term average algorithm and the prediction algorithm occur energy outage on the 6th day when the BS has few energy arrivals. Even though the conservative algorithm doesn’t suffer energy outage, it’s energy output has the most violent temporal variance, which results in its low utility.

In Fig. 8, battery capacity is changed from 6 hours to 48 hours. When the battery capacity is lower than 18 hours, three heuristics have negative utility. Our prediction and optimization algorithm achieves nearly optimal performance throughout the battery capacity. When the battery capacity is large enough, the offline algorithm and our prediction and optimization algorithm have average utility as 0.1764 and 0.1404, respectively, which is quite close. This figure also
shows that in order to achieve the best utility, the battery should be able to store more than 1.5 days’ energy arrivals.

VI. CONCLUSIONS

This paper shows our preliminary work on the solar energy modeling and BS energy allocation in cellular networks based on 27-year solar radiation data. We present the statistic characteristics of solar energy arrivals and give a simple and general way to predict future’s solar energy arrivals. To evaluate the performance of different algorithms, we assume a greedy source data traffic and make use of a utility function. Based on simple assumptions, simulation results show that our proposed prediction and optimization algorithm greatly improves the utility compared with other online heuristics. Future work may include the solar energy allocation problem considering both solar energy prediction and data traffic model.

ACKNOWLEDGMENT

This work is sponsored in part by the National Basic Research Program of China (2012CB316001), the Nature Science Foundation of China (61201191, 60925002, 61021001) and Hitachi R&D Headquarter.

REFERENCES