

GM-PAB: A Grid-based Energy Saving Scheme with Predicted Traffic Load Guidance for Cellular Networks

Rongpeng Li^{*†}, Zhifeng Zhao^{*†}, Yan Wei^{*†}, Xuan Zhou^{*†}, and Honggang Zhang^{*†}

^{*}York-Zhejiang Lab for Cognitive Radio and Green Communications

[†]Dept. of Information Science and Electronic Engineering

Zhejiang University, Zheda Road 38, Hangzhou 310027, China

Email: {lirongpeng, zhaozf, weiyang1987, zhouxuan, honggangzhang}@zju.edu.cn

Abstract—In cellular networks, the base station power consumption is not simply proportional to the traffic loads of its coverage. As the traffic load fluctuates spatially and temporally, the base stations consequently suffer from heavy energy wastage when the traffic loads of their coverage are low. In this paper, we propose a grid-based energy saving scheme over predicted traffic loads. We firstly take advantage of the spatial-temporal pattern of traffic loads and employ the compressed sensing method to predict the future traffic loads. Then, we propose a grid-based energy saving scheme to improve the energy efficiency through turning some base stations into sleeping mode while ensuring the quality of service. Results of the simulation with real traffic loads¹ finally show the accuracy of the traffic load prediction and large energy efficiency improvement.

I. INTRODUCTION

With the increasing popularity of iPhone and other smartphones, we are experiencing an explosive demand for radio network access, while it is incurring large power consumption[1]. Furthermore, it is coming to a consensus that the information and communication technologies (ICT) industry has emerged as one of the major contributors to the world power consumption and green house gas emission. Recent studies show that the ICT industry accounts for 2% to 10% of the world's overall power consumption [2] and it's envisioned that the energy exhaustion² of mobile networks would reach 124.48 kWh in the year of 2011[3] and the power bill will doubly enlarge in five years for China Mobile[4]. Beyond that, increasing awareness of potential environmental impact induced by greenhouse gas emissions and exhaustion of non-renewable energy resources also arise the need to improve the energy efficiency of the telecommunication systems.

In addition to the environmental concerns, there are also economical benefits for cellular network operators to reduce the power consumption of their networks. The energy expenditure accounts for a significant proportion of the overall cost and great savings in capital and operating expenditure

¹It's worthwhile to note here that we hypothesize some power consumption parameters of base stations on the basis of practical scenarios.

²In this paper, we deem the terminologies energy exhaustion and power consumption are equivalent.

(CAPEX and OPEX) can be achieved through improved energy efficiency.

Currently, the subscribers cost around 1% of the overall energy in ICT industry. Comparatively, over 80% of the power consumption takes place in the radio access network, especially the base stations (BSs)[5]. The reasons may lie in the networks at present stage are designed for peak throughput and deployment of BSs is usually optimized for maximum user demand operation. However, there are significant variations in traffic loads, both temporally and spatially[6] since that the typical day-night division of user behavior causes the temporal periodic pattern, and the spatial difference occurs when people move between their apartments and downtown districts back and forth, taking along their mobile terminals. Unfortunately, present communication infrastructure's activities in reality are only weakly dependent on the amount of the traffic loads, which means certain BSs still have to undergo air conditioning or cooling and monitoring and consume relatively large amount of energy even when there is little user accessing requirement, resulting in severe energy wastage[7]. Therefore, there exists the possibility to improve the energy efficiency of base stations by adapting the working status of base stations to the traffic loads.

In a nutshell, this paper considers how to optimize the number of active base stations, depending on the predicted traffic loads. The contributions of this paper lie in that we employ the spatial-temporal characteristics of traffic loads and adopt a compressed sensing based method to predict the future traffic loads. Given the predicted traffic loads, instead of merely minimizing the number of active base stations, we propose a grid-based scheme to minimize the power consumption of active base stations (GM-PAB) by balancing the number of active base stations and their coverage (and also their served traffic loads). Beyonds that, in order to give an accurate assessment for our research, we simulate with real traffic loads in Hangzhou, China and prove large energy efficiency improvement.

The remainder of the paper is organized as follows. In Section II, we cover the traffic load prediction model and introduce the background and power consumption model of base

stations. In Section III, we talk about the compressed sensing-based traffic load prediction method. Section IV introduces the grid-based energy saving scheme. Section V evaluates our proposed schemes and represents our simulation results. Section VI concludes this paper with a summary.

II. SYSTEM MODEL

A. Background on Base Stations

The radio access network of cellular networks is responsible for handling traffic and signaling. In the typical architecture of radio access network in cellular networks, base station controllers (BSC) and their controlled tens of base stations are the main part of the radio access network.

[8] shows the power consumption for one base station can be approximated as $P_t = P_\alpha \cdot V + P_\beta$ with respect to the traffic loads V . In this paper, we deem P_α as a constant in certain transmission range, but it varies over the transmission distance since the base station will exhaust more energy to serve the traffic loads from longer distance. The constant part P_β is due to the cooling system and other auxiliary devices owing to that their power consumption mainly depends on the stationary working environment. It's worth noting here that the supporting subsystem contributes to the overall power consumption as high as 50% [5]. Hence, cellular networks can save a large amount of energy if base stations can enter the sleeping mode when few traffic loads exist.

B. Traffic Load Prediction Model

In this paper, we assume that there exists an n -length traffic load vector to timely record the volumes of one BSC controlled n base stations' traffic loads at one specific moment. For monitoring purposes, the volumes of traffic loads should be traced periodically to better know the status of base stations. Hence, we refer to traffic load matrix as the set of traffic load vectors at different moments. For example, $X(i, t)$ in a traffic load matrix X denotes the volume of traffic loads of base station i at the moment of t . Therefore, the traffic load prediction problem can be formulated as determining the traffic load vector in the near future moment τ . Unfortunately, since some traffic volumes cannot be recorded due to the possible overloading of recoding servers, the prediction issue has to be based on the incomplete or inaccurate measured traffic load matrix.

III. COMPRESSED SENSING-BASED TRAFFIC LOAD PREDICTION

In real world, a great many of signals or datasets exhibit characteristics like structure or redundancy. To be more specific, the representation of the signals or datasets has only a small number of non-zero elements in some transformation bases. Compressed sensing, which aims to handle these sparse signals problems, has developed and attracted large attention recently[9]. According to the compressed sensing theory, if the signals match the sparsity condition, we can make use of this prior knowledge for both sampling and reconstruction of the signals or datasets with much smaller-dimensional

measurements. Meanwhile, compressed sensing or its usually embedded greedy algorithm is also an alternative way to solve under-constrained linear inverse problems.

In the context of matrix compressed sensing, the matrices which have very low rank will have the sparsity, since the vector of singular values generated from a low rank matrix is sparse under the constraint. As the traffic load matrix has certain periodic pattern[6], the traffic load matrix can be approximated as a low-rank matrix. So, the compressed sensing can be applied in the traffic load prediction problem.

We firstly concentrate on how to find a low-rank matrix \hat{X} , whose entries are almost equal to those in X , except the missing or inaccurate values in the latter matrix X . We adopt a factorization method to approximate the incomplete matrix X , while the factorization result \hat{X} has the least rank and meanwhile satisfies the following equation

$$\begin{aligned} \min \quad & \text{rank}(\hat{X} = LR^T) \\ \text{s.t.} \quad & M * (LR^T) = M * X, \end{aligned} \quad (1)$$

where M is a $m \times n$ matrix and the entries of M are 0 when the corresponding entry in X is missing while they are 1 otherwise. $M * X$ is the element-wise multiplication. Moreover, the sizes of factorization matrix L and R are $m \times r$ and $n \times r$ respectively and r is a factor taking the value far less than $\min(m, n)$.

When \hat{X} is determined by LR , the missing values in X can be approximated using their counterparts in \hat{X} . Therefore, the only remaining problem is to solve Eq.1. Unfortunately, the objective of rank minimization is not convex and is perhaps not solvable. But under certain particular constraints (e.g., the restricted isometry property in compressed sensing[10]), we can transform the rank minimization in Eq.1 to the following equivalent form:

$$\begin{aligned} \min \quad & \|L\|_F^2 + \|R\|_F^2 \\ \text{s.t.} \quad & M * (LR^T) = M * X, \end{aligned} \quad (2)$$

where the form $\|\bullet\|_F$ denotes the Frobenius norm, with $\|Y\|_F = \sqrt{\sum_{i,j} Y(i,j)^2}$ for a matrix Y .

Besides, due to the relative low rank nature of the original data matrix X and the inaccuracy of the measurements, the matrix factorization that perfectly fits the constraints may not exactly exist. As a result, we include the precision of the approximation as a part of the optimization process. Thus, we can transform Eq.2 into

$$\min \|M * [(LR^T) - X]\|_F^2 + \alpha(\|L\|_F^2 + \|R\|_F^2), \quad (3)$$

where α is the weight factor. In this regard, this approach finds a low rank regularized factorization, which fits the intrinsic characteristics of the data matrix X . In addition, it does not strictly ensure the constraints, but still keeps the accuracy by minimizing the difference.

Actually, we can utilize the spatial and temporal relevancies in the original datasets to make the approximation more precise[11]. According to the empirical measurements, in temporal dimension, the traffic values in adjacent moments are

often close. Also, in spatial dimension, the traffic values between neighboring places also have some kind of smoothness due to the space relevancy[6]. Therefore, we can utilize this prior knowledge to reconstruct the traffic load matrix, which can be formulated as follows

$$\min \|M.*[(LR^T) - X]\|_F^2 + \alpha(\|L\|_F^2 + \|R\|_F^2) + \beta\|SLR^T\|_F^2 + \gamma\|LR^T T^T\|_F^2, \quad (4)$$

where T and S denote the temporal relevancy matrix and spatial relevancy matrix, respectively. β and γ are weight factors like α as well.

We have stated the way to determine the missing values in the incomplete traffic load matrix by the compressed sensing-based method. Indeed, if we augment the known incomplete traffic load matrix with the traffic load vector at the future moment τ , we can obtain the future traffic load vector with the compressed sensing-based algorithm above. As to temporal relevancy matrix, we refer to [6]'s design insight 2: *at any time, the traffic load difference in two consecutive days is less than 20% for 70% base stations*. Therefore, we can construct a temporal relevancy matrix as

$$T = \begin{matrix} & \overbrace{\hspace{10em}}^{24hours} & \\ \begin{bmatrix} 1 & 0 & \cdots & -1 & \cdots & \cdots & \cdots \\ 0 & 1 & 0 & \cdots & -1 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots & -1 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \end{bmatrix} & \end{matrix}. \quad (5)$$

Then, using the spatial-temporal compressive sensing method with the temporal relevancy matrix, we are able to predict the traffic loads of future time based on the historical traffic loads and underlying trends.

IV. TRAFFIC PREDICTION-BASED BASE STATION ENERGY SAVING SCHEME

Currently in typical cellular networks, quite a lot of base stations are under utilization at most time of a day, which results in much energy wastage and heavy energy inefficiency. An intuitive solution[12] is to minimize the number of active base stations (M-NAB) when some of base stations has very low traffic loads. The M-NAB solution gives the approach for base station energy saving, but there are several drawbacks in it. First of all, it does not concern how to adjust the base stations' coverage to transfer or reallocate the traffic load. Secondly, the M-NAB solution does not consider the increased power consumption for one base station to zoom its coverage or serve larger traffic loads. Owing to the drawbacks of M-NAB, we propose a grid-based base station energy saving scheme.

In the first place, we divide the coverage of every base stations as certain grids with equal size. Because of that the coverage of one base station is usually small, we assume the traffic loads of one base station is equally distributed in every grid in its coverage. Then we map the base stations and grids as the vertices of an undirected graph $G = (V, E)$, which are represented as *triangles* and *squares* respectively for the

convenience of representation. There exists an edge between the grid (*square*) and one of its surrounding base stations (*triangle*) so long a this grid is located in the base station's maximum transmission range. Thus, the N-MAB solution can be regarded to select the edges to form a graph with minimal number of edges while keeping in mind that the *square* need connect with one *triangle*. Additionally, the *triangle* should be connected with limited *squares*, ensuring the base station can provide so many traffic loads. So, we can formulate the description as (we call it as Grid-based M-NAB, or GM-NAB):

$$\begin{aligned} \min & \sum_{i=1}^n \text{sgn}(\sum_{e_{i,k} \in E} I_{e_{i,k}}) \\ \text{s.t.} & \sum_{e_{i,k} \in E} V_{e_{i,k}} \times I_{e_{i,k}} \leq C_i, \quad \forall i \in 1 \cdots n \\ & \sum_{e_{i,k} \in E} I_{e_{i,k}} \geq 1, \quad \forall k \\ & I_{e_{i,k}} \in \{0, 1\}, \quad \forall e_{i,k} \in E \end{aligned} \quad (6)$$

where $e_{i,k}$ denotes the edge between base station i and grid k while $V_{e_{i,k}}$ denote the traffic loads in grid k . $I_{e_{i,k}} = 1$ indicates $e_{i,k}$ is included in the optimal solution while $I_{e_{i,k}} = 0$ indicates not. C_i represents the capacity threshold of base station i . The constraints guarantee the traffic loads distributed to any base station do not exceed its capacity and every grid area can be covered by one base station.

The GM-NAB solution helps explain how to adjust the coverage of base stations, yet it does not take into consideration that some proportion of the power consumption is relevant to the traffic loads and transmission distance. To solve that drawback, we introduce the edge weight $P_{e_{i,k}}$ in graph $G = (V, E)$ to denote the power consumption for base station i to serve the traffic loads in grid k . The value of edge weight $P_{e_{i,k}}$ depends on the traffic loads and distance between base station i and grid k . The longer transmission distance is, the larger edge weight $P_{e_{i,k}}$ will be. Moreover, let $P_{i,steady}$ denote the stationary part of power consumption for base station i . Hence, we can extend the objective function in Eq.6 to the grid-based minimization of the power consumption for active base stations (GM-PAB), namely

$$\min \sum_{e_{i,k} \in E} P_{e_{i,k}} \times I_{e_{i,k}} + \sum_{i=1}^n P_{i,steady} \times \text{sgn}(\sum_{e_{i,k} \in E} I_{e_{i,k}}), \quad (7)$$

As there is a sign function in Eq.7, whose discontinuation makes the equation difficult to solve, we make some transformation on Eq.7

$$\begin{aligned} \min & \sum_{e_{i,k} \in E} P_{e_{i,k}} \times I_{e_{i,k}} + \sum_{i=1}^n P_{i,steady} \times I_{s_i} \\ \text{s.t.} & \sum_{e_{i,k} \in E} V_{e_{i,k}} \times I_{e_{i,k}} \leq C_i, \quad \forall i \in 1 \cdots n \\ & \sum_{e_{i,k} \in E} I_{e_{i,k}} \geq 1, \quad \forall k \\ & \sum_{e_{i,k} \in E} I_{e_{i,k}} - I_{s_i} \times N_{s_i} \leq 0, \quad \forall i \in 1 \cdots n \\ & I_{e_{i,k}} \in \{0, 1\}, \quad \forall e_{i,k} \in E \\ & I_{s_i} \in \{0, 1\}, \quad \forall i \in 1 \cdots n \end{aligned} \quad (8)$$

where $I_{s_i} = 1$ indicates base station i will remain active in the final energy saving scheme while $I_{s_i} = 0$ indicates not.

And N_{s_i} is the number of edges connected to base station i 's corresponding triangle in graph $G = (V, E)$. The newly added constraint ensures that when base station i is selected to enter into sleeping mode ($I_{s_i} = 0$), there is no need for its service to any grid (all the corresponding $I_{e_i,k} = 0$).

The representation in Eq.8 for GM-PAB solution is a typical binary integer programming problem, which is usually an *NP-hard* problem. Fortunately, there is a vast of mathematician algorithms by providing good approximation solutions. In this paper, we solve our GM-PAB problem by the integration of branch-and-bound algorithm[13] and primal and dual algorithm[14] embedded in Mosek Optimization Tools[15].

V. EVALUATION AND ANALYSIS

A. Performance of Traffic Load Prediction Method

To evaluate the performance of compressed sensing-based traffic load prediction algorithm, we intentionally hide some existing data, and then apply the prediction algorithm. Thus, this methodology provides us with ground truth for comparison so long as the pseudo-unknown data is not used in prediction algorithm in any way.

To better assess the performance of compressed sensing-based prediction algorithm, we run two independent simulations over one two-week traffic load database in Hangzhou. Both of the two simulations utilize six-day training database to predict the next-day traffic loads. Then, we measure the prediction accuracy by the normalized root mean square error (NRMSE).

As Fig.1(a) shows, the NRMSE for these two independent simulations is less than 0.3 at most of the time. Fig.1(b) also shows the approximation between the predicted traffic loads and original data. In the whole, the performance of the introduced prediction algorithm is good and stable under different traffic load situations.

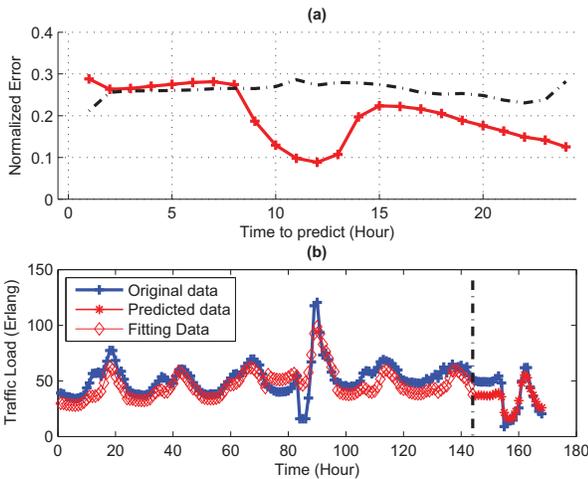


Fig. 1. Performance of traffic load prediction using spatial-temporal compressed sensing method: (a) NRMSE for two independent simulations; (b) prediction result for one selected base station in one simulation.

B. Performance of Grid-based Base Station Energy Saving Solution

To verify the performance of our grid-based solution, we consider 64 base stations with equal size of coverage (1km in diameter). And the main parameters in our simulation are heterogeneous and assumed as follows: (1) the capacity of each base station is 110% of the maximum traffic loads for a given base station; (2) the maximum transmission range also varies in base stations and can be 1.6km and 2km consistent with many available products[6]; (3) the constant part of the power consumption P_β can be 2100W and 2800W; (4) P_α differs in 4, 6 and 8 as the transmission distance varies in 1km, 1.6km and 2km, depending on the base station capacity, we assume one third of base stations' constant power consumption is 2800W with 2km's maximal transmission range. Other base stations' constant power consumption is 2100W with 1.6km's maximal transmission range.

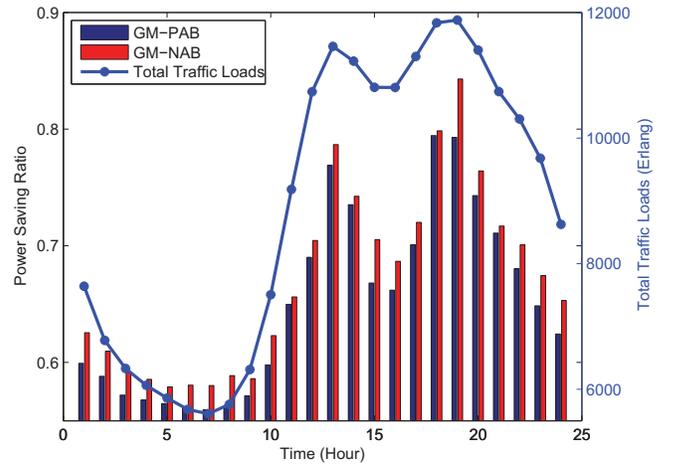


Fig. 2. Power saving ratio of GM-PAB solution and GM-NAB solution in one whole day and the corresponding traffic loads.

We define two measurements, namely power saving ratio and number of active base stations ratio to evaluate the performance of GM-PAB solution and GM-NAB solution. To be specific, we define power saving ratio as the power consumption after utilizing the GM-PAB or GM-NAB solution over the power consumption without any energy saving scheme while number of active base stations ratio is similarly defined. Fig.2 and Fig.3 show how the performance of our energy saving schemes varies at each hour of a day under the scenario each base station's coverage is divided into 64 grids of same size. We can see the performance is consistent with the overall traffic load variations. Fig.3 shows that we can turn half of the base stations into sleeping mode in the earliest 10 hours of a day and save almost 40% of the power consumption during these first 10 hours after taking energy saving schemes as indicated in Fig.2. Beyond that, if we utilize the energy saving schemes, the power consumption of the whole day can be reduced as almost 65% of the consumption before.

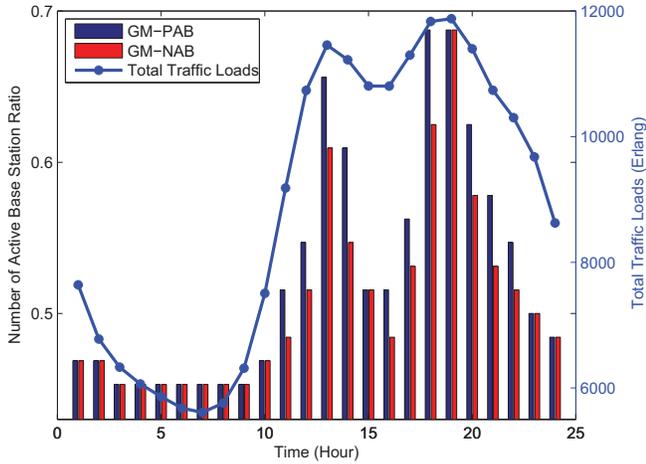


Fig. 3. Number of active base stations ratio of GM-PAB solution and GM-NAB solution in one whole day and the corresponding traffic loads.

Especially, the proposed GM-PAB solution can save 34.6% of the power consumption, roughly 6.3% more energy savings over the GM-NAB solution as noted in Fig.2 even though the number of active BSs for GM-NAB solution is smaller than that for GM-PAB solution if we compare the two histograms in Fig.3. It's the evidence of merely minimizing the number of active BSs can only bring suboptimal results since the active BSs have to zoom out their coverage to serve more traffic loads, leading to higher power consumption for them.

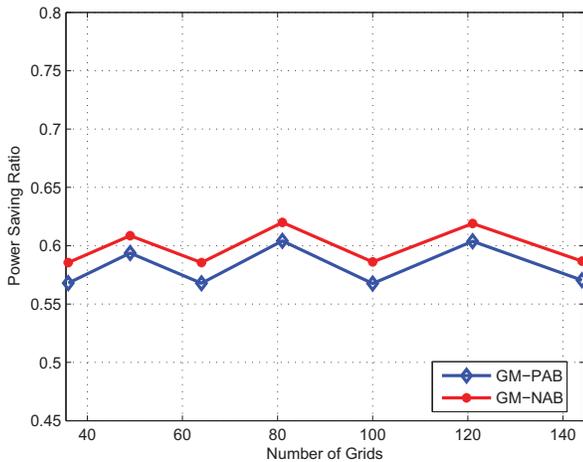


Fig. 4. Performance of GM-PAB solution and GM-NAB solution versus the number of grids in one base station's coverage.

Fig.4 shows how the power saving ratios of energy saving schemes vary as the number of one base station's coverage. We can see that the power saving ratios are always around 58% even though there are small fluctuations as the number of grids in one base station's coverage increases from 36 to 144. Hence, it verifies the effectiveness and robustness to adopt the GM-PAB solution to improve the energy efficiency in cellular

networks.

VI. CONCLUSION

This paper studied how we can save the power consumption of cellular networks by turning some base stations into sleeping mode, based on the predicted traffic loads. We have shown that the traffic loads of cellular networks follow some spatial and temporal patterns, thus making the traffic load matrix satisfy the low-rank property. In that regard, we employ the compressed sensing-based prediction method to forecast the future traffic loads. According to the prediction result of traffic loads, we propose a grid-based energy saving scheme GM-PAB to turn some base stations into sleeping mode while guaranteeing the quality of service and we formulate it as a binary integer programming problem. The simulation results verify the precision of our prediction method and the effectiveness of our energy saving schemes at last.

ACKNOWLEDGMENT

This paper is partially supported by the National Basic Research Program of China (973 Program 2012CB316000) and the National Natural Science Foundation of China (NSFC) under grant number 61071130.

REFERENCES

- [1] H. Zhang, A. Gladisch, M. Pickavet, Z. Tao, and W. Mohr, "Energy efficiency in communications," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 48–49, Nov. 2010.
- [2] M. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal energy savings in cellular networks," in *Proceedings of IEEE ICC 2009*, Dresden, Germany, Jun. 2009.
- [3] R. ABI, "Mobile networks go green-minimizing power consumption and leveraging renewable energy," 2009. [Online]. Available: <http://www.abiresearch.com/>
- [4] C. Institute, "C-ran: Road towards green radio access network," White Paper, V1.0.0, Tech. Rep., 2010.
- [5] G. P. Fettweis and E. Zimmermann, "ICT energy consumption-trends and challenges," in *Proceedings of WPMC 2008*, Lapland, Finland, Sep. 2008.
- [6] C. Peng, S.-B. Lee, S. Lu, H. Luo, and H. Li, "Traffic-driven power savings in operational 3g cellular networks," in *Proceedings of ACM Mobicom 2011*, Las Vegas, Nevada, USA, Sep. 2011.
- [7] L. Herault, E. C. Strinati, D. Zeller, and O. et al, "Green communications: a global environmental challenge," in *Proceedings of WPMC 2009*, Sendai, Japan, Sep. 2009.
- [8] F. Richter, A. J. Fehske, and G. P. Fettweis, "Energy efficiency aspects of base station deployment strategies in cellular networks," in *Proceedings of IEEE VTC 2009 Fall*, Anchorage, Alaska, USA, Sep. 2009.
- [9] D. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 4036–4048, 2006.
- [10] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum rank solutions to linear matrix equations via nuclear norm minimization," 2007, arXiv preprint arXiv:0706.4138.
- [11] Y. Zhang, M. Roughan, W. Willinger, and L. Qiu, "Spatio-temporal compressive sensing and internet traffic matrices," in *Proceedings of ACM SIGCOM 2009*, Barcelona, Spain, Aug. 2008.
- [12] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Communication Magazine*, vol. 48, no. 11, pp. 74–79, Nov. 2010.
- [13] J. W. Chinneck, *Practical optimization: a gentle introduction (Chapter 12)*. [Online]. Available: <http://optlab-server.sce.carleton.ca/POAnimations2007/BalasAddAlg.html>
- [14] S. P. Bradley, A. C. Hax, and T. L. Magnanti, *Applied Mathematical Programming (Chapter 9)*. Addison-Wesley Press, 2007.
- [15] *The MOSEK optimization tools manual*, 2011. [Online]. Available: <http://docs.mosek.com/6.0/tools/index.html?id=2>