Spatial-Temporal Compressed Sensing Based Traffic Prediction in Cellular Networks

Qian Wen*[†], Zhifeng Zhao*[†], Rongpeng Li*[†], and Honggang Zhang*[†]

*York-Zhejiang Lab for Cognitive Radio and Green Communications

[†]Department of Information Science and Electronic Engineering

Zhejiang University, Zheda Road 38, Hangzhou 310027, China

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

Abstract—In conventional cellular networks, base stations (B-Ss) usually suffer from severe power consumption since they are working to guarantee the coverage and QoS (quality-of-service) requirement according to the peak traffic load generated by the mobile cellular users Accordingly, how to precisely forecast the future traffic load to promote network cooperation and adaptive energy resource allocation in complying with the variation of spatial-temporal traffic load has been an emerged issue due to the significant energy exhaustion of BSs. In this paper, we propose a spatial-temporal compressed sensing based network traffic prediction method to solve this problem. We first construct a traffic matrix (TM) by using previously measured data and setting the data to be predicted as zeros, corresponding to the volume of traffic load. Then, compressed sensing approach with large scale and small scale temporal constraints as well as spatial constraints is employed to factorize the traffic matrix. By reuniting the results of traffic matrix factorization, we obtain the estimation of predicted traffic data. Numerical results have showed that this method can restrict the prediction error under 10% when dealing with real traffic load data.

I. Introduction

In current cellular telecommunications systems, network resources are always allocated according to the peak traffic load of mobile terminals (MTs) to guarantee the network coverage and meet the explosive transmission rate demand. However, user traffic flow in cellular mobile networks has exhibited a fundamental feature of uniformity and uncertainty in terms of temporal or spatial variation pattern. Therefore, there exists a period when the traffic flow are relatively low and the base stations are severely under-utilized, leading to huge energy wastage [1]. Recently, more and more intense efforts have been dedicated to solve the energy wastage problem while guaranteeing the communication quality of service (QoS), most of which attempt to adaptively switch on/off some of BSs or dynamically allocate the networks resource in complying with the traffic load characteristics [2]-[4]. Accordingly, how to precisely prefigure the accurate network traffic load at certain time for each BS is becoming a critical issue with ever-growing significance. In the meanwhile, the typical division of day-night living behavior of cellular users generates temporal and spatial periodic fluctuation in the daily traffic load (we also call it "tidal phenomenon" of traffic), which also demonstrates the feasibility to undergo traffic load prognostication and dynamic network resources allocation [5].

Up to now, several theoretical analyses related to traffic

load forecast have been investigated in the scenario of cellular networks. Typically, traffic flow estimation has been assumed as an under-constrained linear-inverse problem either temporally or spatially which can further be solved based on a prior model of the traffic matrix (TM). To be specific, the previous researches firstly approximate the traffic volume with a certain distribution function. For example, it's commonly assumed that users arrive according to a Poisson process while the sojourn time is exponentially distributed [2]. Afterwards, specific models such as Markov processed [6] and regression models [7] are applied to capture the temporal dynamics. As for spatial distribution, log-normal distribution model [8] and exponential distribution model [9] are adopted. Besides, other models also are taken into account. For example, [10] modeled traffic profile in real cellular radio access networks as a periodic sinusoidal function. When we apply these models to predict the real traffic loads, we merely need to choose the proper concrete parameters to fit the models with the existing data. They are simple to use, but the simulated values according to these models are too idealistic while far from precisely matching with the actual traffic load in reality. In the meanwhile, the parameters decision process is very sensitive to the various traffic flow dynamics and physical networking scenes. As a result, we experience large errors when predicting traffic flow using the above methods.

In this paper, we consider a new approach without solely relying upon a deterministic analysis model to estimate the networks traffic load but concentrating on the values of the traffic flow themselves. To begin with, we collect all the traffic data of the BSs in a certain area as a traffic matrix (TM). Due to the temporal and spatial periodic variation of cellular user mobility pattern, there is high redundancy and correlation among the data in the TM. In [11][12], spatial-temporal compressed sensing (STCS) based on singular value decomposition (SVD) is employed to solve the matrix data prediction problem. The state of the art to employ STCS is to find a proper temporal and spatial constraints matrix that can regularize the data prediction process. In this paper, we refine the prediction process in [11] and [12] by considering heuristic temporal constraints from largescale as well as smallscale perspective. Moreover, the networking traffic load among the neighboring BSs may possess some spatial relevance due to the spatial-temporal characteristics of users' behaviors and
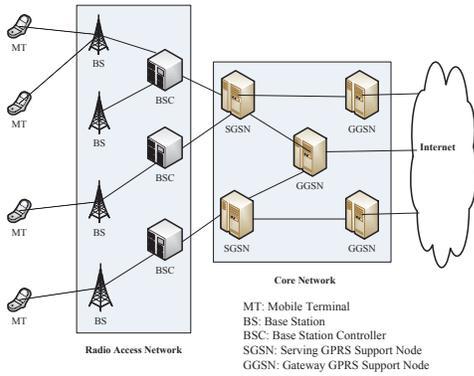
Fig. 1. Architecture of a cellular network.

interests, so we also take the topology and the previous data as a prior knowledge to seek the inner spatial relationship among them to improve the prediction accuracy. The simulations validate that the STCS based prediction method with the proposed temporal and/or spatial constraint refinedness can outperform the methods in [12] with an over 10% prediction accuracy improvement.

The rest of the paper is organized as follows: the system model including physical scene and traffic prediction model formation is covered in Section II. In Section III, we introduce the solution of traffic prediction problem with spatial-temporal compressed sensing. This section also contains the proposed large scale and small scale spatial-temporal constraint matrices. Numerical results are provided in Section IV Section V concludes the paper with a summary.

II. System Model

Fig.1 illustrates the architecture of the cellular network. As Fig. 1 depicts, the cellular network mainly consists of three parts: (1) mobile terminals (MT); (2) radio access network (RAN); (3) core network (CN). The radio access network includes BSs and base station controllers (BSCs), which connect the MTs to CN though the serving GPRS support nodes (SGSNs). In CN, the SGSNs connect to the internet through the gateway GPRS support nodes (GGSN).

A. Traffic Prediction Model Formation

In this paper, we assume that each BS records the volumes of voice service i.e., non-negative values in Erlang, at different specific moments as a traffic load vector and transmit it to its controlling BSC, then CN Moreover, we experiment with the realistic flow level mobile device traffic data collected from the cellular core network. As traffic load is general continuous, we treat the volumes of traffic load in time interval $[t, t + \Delta t']$ as an entry in the vector. Thus, we can obtain an N length vector to indicate the volumes of traffic load in N intervals for a specific BS. Afterwards, we augment M traffic vectors from M BSs to a traffic matrix (TM) X, where $X(i, j)$ denotes the volume of traffic load of BS i at the time interval j. In a nutshell, each row vector indicates the volumes of traffic load for a specific BS and each column vector denotes the volumes

of traffic load for adjacent BSs at a certain interval. Note that, as X is collected from CN, the measurement volumes of traffic load might be burdened with missing or abnormal values. Therefore, traffic load prediction problem can be explicated as finding the explicit values of the traffic load at the future moments $\Delta \tau$ from an incomplete TM X.

Since there might be some abnormal or missing values in the TM because of the recording error or transmission packet loss in networks, the conventional prediction methods might be unattainable in this scene. Hence, we proposed a spatial-temporal compressed sensing based traffic prediction approach in Section III.

III. Spatial-Temporal Compressed Sensing Based Traffic Prediction

A. Spatial-Temporal Compressed sensing

By virtue of the temporal and spatial variance of user mobility, TM contains certain periodic pattern [3] or some structure patterns and redundancy [13]. Previous researches have suggested that typical TM inherits a relatively low-dimensional subspace [13] [14]. Therefore, it's feasible to take the TM as a low-rank matrix. So, in some sense, the prediction problem can be regarded to find a low-rank approximation $\widehat{X}$ whose entries best match the values in X.

Singular value decomposition (SVD) is a basic tool for such a low-rank matrix approximation. It's common to rewrite X as the equivalent form

$$X = U \Sigma V = L R^T \quad (1)$$

where U is an $n \times n$ unitary matrix and V is an $m \times m$ unitary matrix, and Σ is an $n \times m$ diagonal matrix containing the singular values of X. $L = U \Sigma^{1/2}$ and $R = V \Sigma^{1/2}$. A factorization solution could be found by settling the following rank minimization problem [12]

$$\begin{aligned} &minimize \ rank(X = L R^T), \\ &\text{subject to } M.*LR^T = M.*X \end{aligned} \quad (2)$$

where M is an $m \times n$ measurement matrix which to be designed as the linear operator to X. Unfortunately, the objective of the above rank minimization is non-convex and thus difficult to settle. But somehow, a heuristic idea from the compressed sensing works perfectly here [11]. To be specific, if M has the restricted isometry property (RIP) [16] [17], we can transfer the rank minimization as follows

$$\begin{aligned} &minimize \ ||L||_F^2 + ||R||_F^2 \\ &subject \ to \ M.*LR^T = M.*X \end{aligned} \quad (3)$$

where $||*||_F$ is Frobenius norm, with $\|Y\|_F = \sqrt{\sum_{i,j} Y(i,j)^2}$ To meet RIP and simplify this problem, it's feasible to set M as a unit matrix.

However, as the low rank property of the original traffic load matrix and the uncontrolled error in the measurement, the results of (3) might not be the perfect approximation of X. So we use the following optimization objective with the pre-mentioned unit matrix M to replace (3)

$$minimize||(LR^T) - X||_F^2 + \alpha(||L||_F^2 + ||R||_F^2) \quad (4)$$

where α is a parameter that reflects the tradeoff between precisely approximating the measurement data and the ambition of achieving a low rank matrix.

To solve the optimization of (4), we initially choose L and R randomly. Then we alternatively drive (4) with a fixed L or R per round until it reach the convergence.

As mentioned before, there exists strong temporal and spatial periodic division of traffic load. Therefore, it's reasonable to concrete the prior knowledge about spatial-temporal structure of TM, e.g., the correlations among rows and columns, to refine the optimization. So we obtain

$$minimize||(LR^T) - X||_F^2 + \alpha(||L||_F^2 + ||R||_F^2) \\ + \beta||SLR^T|| + \gamma||LR^T T^T|| \quad (5)$$

where S and T denote the spatial constraints matrix and temporal constraints matrix, respectively Similar to α, β and γ are parameters, showing the tradeoff between precisely approximating the measurement data and the target of getting low rank. T is designed to explore the temporal relevance among different time intervals for a specific BS while S intends to dig out the traffic load relationship among adjacent BSs. So by minimizing $||SLR^T|| + ||LR^T T^T||$, the approximation $\widehat{X} = LR^T$ would inherit the intrinsic spatial and temporal characteristics of measurement X.

B. Choice of Temporal Constaints Matrix

Without loss of generality, we assume there exist a TM with the size of $M \times (N_T \times 24 + N_P)$, where $N_T \times 24$ is the training length of traffic data (hour) we already known, and N_p is the prediction length of data (hour) with $N_T \times 24 + N_P = N$. In [13], T =toeplitz(0,1,-1), which just simply considers the similarities between data of TM in adjacent time intervals and treat the optimization as an minimization of the difference of them. However, it's yet accurate enough because of the inadequacy use of the temporal connections of data and ignorance that there might be large dissimilarities between the on-work hour and it's adjacent off-work hour.

To accurately dig out the temporal variance of traffic load, we first research on the general traffic load in a month. From Fig.2, it's obvious the traffic load is quite self-similar between weeks with the data in a similar shape in weekdays and a little lower volume in weekend, which means the traffic load in one week would experience a large scale variance. Moreover, as illustrated in Fig.3, we can conclude that one day's traffic load also has a temporal small scale variance, that is, the traffic load between days in a week has strong relations. So we take temporal constraints from both large scale perspective small scale perspective.

- **Large scale perspective:** as for the large scale variation, we use the traffic load of previous seven days to forecast the traffic in next several days in next week with the TM as (9). So the temporal constraints can be built up as equation (6).
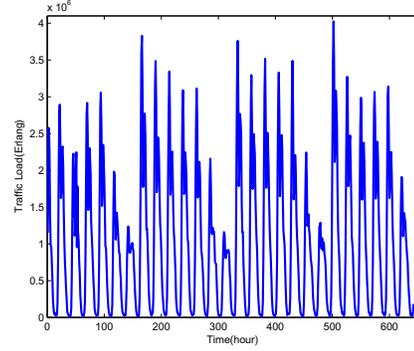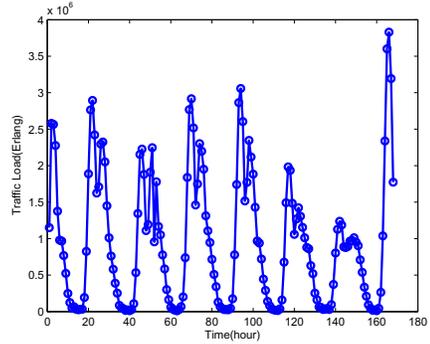

Fig. 2. Traffic load volumes for a BS in a month.

Fig. 3. Traffic load volumes of a BS for a week.

where the weight coefficient in T of the predicted hour at each row is 1 and the others given by the weighted coefficients.$w_l(i)$ is the large scale weight coefficient which can be the average value of the coefficients weighting relations among columns of traffic data of all BSs. It reflects the relevance between the data at the same moment every day with the data to be forecast. Beyond that, we also use w_j, which is the weight coefficients that reflect the relations among the traffic load at the moment to be predicted and previous/next couple of hours to refine the temporal matrix in case the explosively increase of traffic load in some special moment. All this weight coefficients can be determined by exploring the connections in the training traffic data through linear regression.

- **Small scale perspective:** as the small period is one day, we reconstruct the TM into 24 sub TM matrices by choosing the traffic data at the same hour for each day to build up a $(N_T + n_t) \times M$ matrix X_t, where n_t denotes the length of time (hour) to be predicted in the tth sub TM matrix with $\sum_t n_t = N_P$ and $X_t(i,j)$ is the traffic load of ith BS at the tth hour of jth day. So the temporal constraints matrix is built up as equation (7).
where $w_{s,t}(i)$ is a small scale temporal coefficient weighting the relevance among the traffic of prediction moment and previous data. It can be determined by applying

121

$$T = \begin{bmatrix} \overbrace{w_l(1) \quad 0 \quad \cdots \quad 0}^{24hours} & w_l(2) & 0 & \cdots & w_l(N_T) & \cdots & w_1 & 1 & w_2 & \cdots & \cdots & \cdots \\ 0 \quad w_l(1) \quad 0 \quad \cdots & 0 & w_l(2) & 0 & \cdots & w_l(N_T) & \cdots & w_1 & 1 & w_2 & \cdots & \cdots \\ 0 \quad 0 \quad w_l(1) \quad 0 & \cdots & 0 & w_l(2) & 0 & \cdots & w_l(N_T) & \cdots & w_1 & 1 & w_2 & \cdots \\ \vdots \quad \ddots \quad \ddots \quad \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \end{bmatrix} \quad (6)$$

$$T_t = \begin{bmatrix} w_{s,t}(1) & w_{s,t}(2) & \cdots & w_{s,t}(N_T) & 1 & 0 & \cdots & \cdots \\ w_{s,t}(1) & w_{s,t}(2) & \cdots & w_{s,t}(N_T) & 0 & 1 & \cdots & \cdots \\ w_{s,t}(1) & w_{s,t}(2) & \cdots & w_{s,t}(N_T) & 0 & 0 & \cdots & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \end{bmatrix} \quad (7)$$

processing linear regression into the previous data and reach a consensus of the values for a long time.

C. Choice of Spatial Contraint Matrix

Somewhat different from the traffic data of other networks, traffic data in cellular networks has strong local spatial structure because of the periodic user mobility and the similar user interests in certain places. Therefore, the BSs at the neighboring locations with similar interests are burdened with the similar traffic load at the same time. In the prior work [13], k-nearest node (KNN) algorithm, in which the traffic data of k nearest node is averaged as the data of the forecast BS, is used to tackle the spatial problem. Unfortunately it's quite challenging to find the k-nearest nodes of a specific node and then to decide the weights. Besides, using fixed number of nearest nodes is not so flexible when facing various locations of BSs. In this paper, we modify the conventional KNN method by taking advantage of the prior knowledge of the BS deployments and the training data from the graph perspective.

To begin with, notice that, we prefer local spatial constraints rather than global ones, that is, we only considered the spatial constraints among traffic data of BSs that has strong connections. To achieve this, we first figure out the normalized inner product between traffic vectors of every two BSs, namely rows of TM. Afterwards, by assigning a proper threshold λ under the consideration of graph size and values of the whole normalized inner product coefficients, we put BS nodes with the normalized inner product coefficients larger than the threshold into a group to form a vertex set. Once the vertex sets $\{V_i\}$ are determined, every vertex set can be united as a directed connected graph with weighted edges $\{w_{i,k}(j)\}$, where $w_{i,k}(j)$ denotes the edge weights of kth BS in i^{th} graph. The edge weights here indicate the spatial connections of traffic vectors of BSs in a graph. Notice that, we are trying to figure out the long-term average spatial connections among traffic data of BSs rather than correlations at a specific moment. So we apply linear regression algorithm to the previous data for a long period in TM to find a set of weights $w_{i,k}(j)$ so that the linear combination of traffic load

of the BSs in the same graph can best approximate the traffic of kth BSs as follows

$$X(k,*) = \sum_{j=1,j\neq k}^{|V_i|} w_{i,k}(j) \times X(j,*) \quad (8)$$

Traffic load volume of a BS for a week.
Then we assign $\{w_{i,k}\}$ into $S(k,*)$ as follows

$$S(k,j) = \begin{cases} 1, & j = k \\ -w_{i,k}(j), & j \neq k, j \in V_i \\ 0, & other \end{cases} \quad (9)$$

We repeat this process until the whole vertex sets/graphs are traversed and S is built up.

D. Error Constraints

With a fixed length of training data, the prediction error would be in proportion to the length of data to be forecast. So the length of data to be predicted should be limited in a range. This range can be determined by iteratively processing the simulations of current approach with a preset error upper limit λ

IV. NUMERICAL RESULTS

In this section, we present the simulations under various configurations to evaluate the performance of our proposed traffic prediction method with temporal and spatial constraints. To better assess our proposed scheme, we experiment with the realistic traffic data, which contains the traffic loads of 73 BSs in the urban of a typical medium-size city in southeast area of China by China Mobile. As the traffic is continuous in time domain, we calculate the sum of recorded data every one hour as an entry in TM. Besides, we set main parameters for our simulations as TABLE I.

To measure the performance of the approach, we adopt the Normalized Mean Absolute Error (NMAE) with respect to the predicted data and origin ones as equation (10).

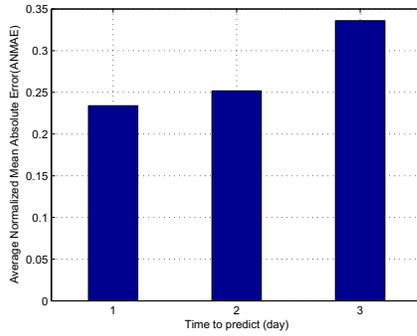$$NMAE = \frac{\sum_{i,j} |X(i,j) - \widehat{X}(i,j)|}{\sum_{i,j} X(i,j)} \quad (10)$$

Fig. 4. Average NAME with different prediction length using temporal compressed sensing based traffic prediction.

Fig. 5. Performance of traffic load prediction using spatial-temporal compressed sensing based methods with different T and S.

Fig. 6. NAME of traffic load prediction based on spatial-temporal compressed sensing with different T and S.

Fig.4 depicts the average error only using temporal constraints to regularize the optimization. To limit the average error around 20% in the time considering the temporal period of traffic load, we use the traffic data of previous 7 days as training data to predict the traffic load of the 8th day. So the TM here is a 73×192 matrix with the 168 columns is training data while the last 24 columns of data are all zeros which are to be forecast

Three methods are simulated for comparison:

TABLE I
SIMULATION PARAMETERS

α	3
β	10^{-3}
γ	10^{-3}
λ	0.2
$MaxIter$	100
N_T(day)	7
N_P(hour)	24

1) KNN method (CS-TK). This method is used in [9] to recover the matrix with missing values. Here we take it as a comparison to our proposed methods. Temporal matrix is designed as (4).

2) Prediction with large scale T and S constraints (CS-LTS). CS-LTS designs the large scale temporal and spatial constraint matrices as (4) and (9) to regularize the prediction, respectively

3) Prediction with small scale T and S constraints (CS-STS). We factorize the whole TM into 24 sub matrix with per matrix predicted the data of an hour in CS-STS Moreover, temporal and spatial constrains is determined as (5) and (9).

In the simulations, we process 10 rounds of simulations and attain the average value as the results.

Fig.5 compares the estimated traffic load with the actual traffic load of a random selected BS with the traffic flow estimated using the above three methods. This result indicates that our proposed approach outperforms the conventional technique. We evaluate the performance of the proposed methods and give the prediction error comparison in Fig.6. While the
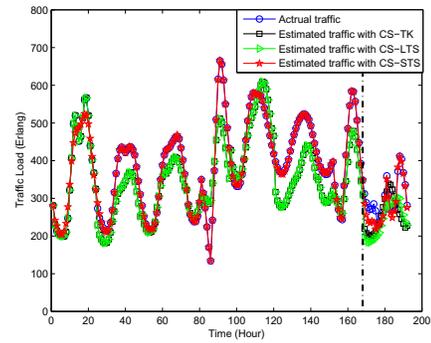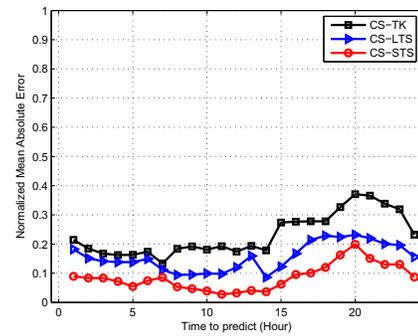
NMAE of the CS-TK mostly lies between 18%-40%, the CS-LTS approach outperforms the CS-TK method with NMAE mostly ranging from 10 to 20%. The reason may result from that in CS-TK method, the number of traffic vectors used to make a linear combination to approximate the traffic vector of the central BSs is fixed while in CS-LTS, it's flexible with respect to the graph size. So CS-LTS is much more adaptive to the training data in TM as to achieve a better performance. It's worthwhile to note here that the NMAE using the CS-STS approach is below 10%. It's quite comprehensive since the traffic values have much more strong connections to the data at the same time in a day. So CS-STS method outperforms CS-LTS method when using these tightly related values to predict the future traffic data. Note that if we only use the spatial constraints to regularize the estimation, the NAME would be too inferior. That is because spatial constraints aim at refining the relationship among the rows in TM whose last few columns are zeros, so any operation of this volumes alone would be useless. Fig.7 explores the impact of the length of data to be forecasted on the prediction performance. We can see that as the average NMAE with the prediction length varying from 1 to 24, the prediction error almost linearly increases.
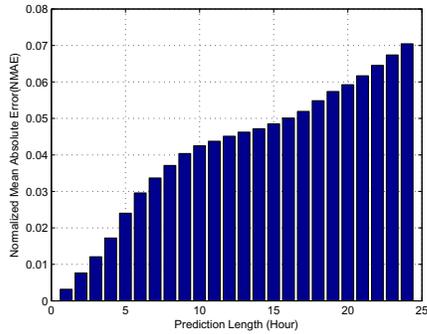
Fig. 7. NMAE of spatial-temporal compressed sensing based traffic load prediction using small scale T and S with respect to prediction length.

V. CONCLUSION

As green communication has emerged as a global issue, precisely traffic flow prediction plays a crucial role in order to enhance BSs cooperation and resources allocation accordingly, thus saving energy consumption while guaranteeing the coverage and QoE of users. However, current traffic forecast methods mainly focus on finding ideal models to simulate the actual traffic dynamics. So the simulated values are often an average or similar values from the long term but far from real temporary ones. In this paper, we transform traffic prediction problem into traffic matrix recovery which can be solved by applying spatial-temporal compressed sensing. Observed that users in cellular network have periodic mobility, we try to dig out the temporal and spatial periodic variation properties in large scale and small scale and refine the corresponding spatial/temporal constraints matrix. Simulations have validated the practicability of our proposed approach with the error ratio less than 10% when using the prior 7 days data to predict the traffic of the following day.

Further research should be explored in the follow aspects. First, improvement of our approach to achieve higher accuracy should be extended. Then a large amount of research should be explored in BS cooperation according to the forecast traffic load to achieve energy efficient network. Besides, different network operation scheme may require discrepant prediction errors. So it is significant to find out the relationship between prediction error constraints and BSs cooperation schemes.

REFERENCES

[1] Honggang Zhang, A. Gladisch, M. Pickavet, Z. Tao, and W. Mohr. "Energy efficiency in communications," *IEEE Communications Magazine*, 48(11):48–49, Nov 2010.

[2] Z Niu, Y Wu, J Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Communication Magazine*, 48(11):74–79, Nov 2010.

[3] C. Peng, SB Lee, S. Lu, H. Luo, and H. Li, "Traffic-driven power savings in operational 3G cellular networks," *Proceedings of ACM Mobicom 2011*, Las Vegas,Nevada, USA, 2011.

[4] Rongpeng Li, Zhifeng Zhao, Yan Wei, Xuan Zhou, and Honggang Zhang, "GM-PAB: A Grid-based Energy Saving Scheme with Predicted Traffic Load Guidance for Cellular Networks," *IEEE ICC 2012*, Ottawa, Canada, June 2012.

[5] Y. Wei, Z. Zhao, and H. Zhang, "Dynamic energy savings in heterogeneous cellular networks based on traffic prediction using compressive sensing" *Proceedings of IEEE ISCIT2011*, Hangzhou, China, 2011.

[6] M. Shafiq, L. Ji, A. X. Liu, and J. Wang, "Characterizing and modeling internet traffic dynamics of cellular devices," *Proceedings of the ACM SIGMETRICS Joint International Conference on Measurement and Modeling of Computer Systems*, San Jose, California, USA, June 0711, 2011

[7] A. Adas, "Traffic models in broadband networks," *IEEE Communications Magazine* vol.35, no.7, pp.82-89, Jul 1997.

[8] Byungchan Ahn, Hyunsoo Yoon, Jung Wan Cho, "A design of macro-micro CDMA cellular overlays in the existing big urban areas," *IEEE Journal on Selected Areas in Communications* (JSAC), vol.19, no.10, pp.2094-2104, Oct 2001.

[9] Ding Zhe and Yu Bin Xu, "A novel spatial Dynamic Spectrum Allocation in the urban heterogeneous scenario," *Proceedings of 2010 IEEE International Conference on Wireless Communications, Networking and Information Security (WCNIS)*, pp.387-391, 25-27 June 2010.

[10] Eunsung Oh and B. Krishnamachari, "Energy Savings through Dynamic Base Station Switching in Cellular Wireless Access Networks," *Proceedings of 2010 IEEE Global Telecommunications Conference* (Globecom 2010), pp. 1-5, Dec. 2010.

[11] D Donoho, "Compressed Sensing" *IEEE Trans. on Information Theory*, 52(12):5406-5425, 2006.

[12] Y. Zhang, M.Roughan, W.Willinger, and L.Qiu, "Spatio-temporal compressive sensing and Internet traffic matrices," *in Proceedings of ACM SIGCOMM* 2009 pp. 267-278, August 2009.

[13] A Lakhina, K. Papagiannaki, M. Crovella, C. Diot, ED Kolaczyk, and N. Taft N, "Structural analysis of network traffic flows," *Procedings of ACM SGIMETRICS 2004*, New York, USA, 2004.

[14] A.Lakhina, M.Crovella, and C. Diot, "Diagnosing network-wide traffic anomalies," *Proc. of ACM SIGMETRICS 2004*, New York, USA, 2004.

[15] Raghu Meka, Prateek Jain, and Inderjit S Dhillon, "Guaranteed Rank Minimization via Singular Value Projection," arXiv:0909.5457v3, 2009.

[16] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," Comptes Rendus Mathematique, 346 (9-10) , pp. 589-592, 2008.

[17] Benjamin Recht Maryam Fazel Pablo A. Parrilo, "Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization," *SIAM Review*, vol. 52, Issue 3, pp. 471-501, 2010.