ITC19/ Performance Challenges for Efficient Next Generation Networks 1701-1710
LIANG X.J. and XIN Z.H.(Editors)
V.B. IVERSEN and KUO G.S.(Editors)
Beijing University of Posts and Telecommunications Press

Spatial Autoregressive Models for Resource Demand Prediction in Mobile Wireless Networks*

Jiazhen Zhou [1], Ken Mitchell [1]

[1] Department of Computer Science & Electrical Engineering,
School of Computing and Engineering,
University of Missouri–Kansas City,
Kansas City, MO 64110 USA.
jzmpc@umkc.edu, mitchellke@umkc.edu

Abstract. Mobile wireless networks present unique challenges to the development of tractable models that can incorporate both spatial and temporal correlations in demand induced by subscriber mobility. Space-time autoregressive time series modeling is a promising inductive method that uses a small number of parameters and can be used for online monitoring and load prediction. In this paper, we develop space-time autoregressive models for several wireless network scenarios. We evaluate the ability of the space-time autoregressive models to model the spatial and temporal correlations in the network and show that for the scenarios depicted, the space-time models perform well.
Keywords: Mobile Wireless Networks, Mobility Prediction, Space-Time ARIMA

1 Introduction

The recent rapid growth in the number of wireless applications, along with the expectation that wireless networks support the same high quality of service provided by wireline networks, present major challenges to service providers who must support these high quality services in a seamless fashion at a reasonable cost. Software radio introduces the ability to dynamically control the wireless interface and is expected to have a huge impact on performance. Ensuring QoS will require innovative solutions with respect to the development of new architectures and protocols, and will provide unique challenges in developing tractable models that can incorporate both spatial and temporal correlations in demand induced by subscriber mobility.

The cellular architecture has been developed to maximize spectrum utilization. Early analytic models focused on the behavior of a single cell [1–3] based on assumptions that arrival traffic was Poisson and homogeneous, mobility patterns were random, and cell sojourn times were exponentially distributed. Models developed under these assumptions provided a tractable analysis and produced reasonably accurate results for first generation wireless networks.

Since there exists no product form solution for multiple cell topologies and the resulting state space explosion prohibits an exact analysis of the network [4, 5], Kelly [4] proposed a fixed-point approximation (FPA) based on the Erlang blocking formula. This method provided a good approximation for blocking probabilities in large networks consisting of multiple cells with low subscriber mobility [6], but has been shown not to be accurate when spatial correlations are strong [7–9].

There has been significant research activity in the utilization of the temporal and spatial characteristics of the physical medium to improve network capacity [10, 11]. In this paper we take on a more global

* This work has been supported in part by US NSF under grant No. ANI 0106640.

approach and concentrate on the modeling and prediction of bandwidth demand as a result of subscriber mobility that is spatially and temporally correlated over a finite time horizon.

To ensure QoS, a number of schemes, most based on priority and admission control, have been proposed. These models typically describe the steady state of the network and enact congestion control or admission policies based on the current state of the network. In this paper we propose a method of predicting the future network state from which we can formulate network management policies.

Autoregressive time series analysis is a powerful inductive modeling tool used to forecast resource demand based on real-time measurements over a finite time horizon. Moreover, these models can readily be identified based on empirical measurements using Kalman filtering [12], least squares methods, maximum likelihood estimates (MLE), or through the use of artificial neural networks [13]. However, autoregressive time series analysis models of networks are generally limited to temporal models [14, 15].

In [16, 17], Pfeifer and Deutsch present a family of multi-variate autoregressive moving average models called space-time autoregressive integrated moving average models (STARIMA). Based on the work of Martin and Oeppen [18], they can capture both temporal and spatial relationships in systems. They have been shown to be a powerful tool in developing parsimonious prediction models in a number of diverse applications from predicting river levels to automobile traffic modeling [16, 17, 19–21].

With the general deployment of 3G wireless systems, empirical data, both temporal and spatial, are becoming available. In this paper we conduct a number of controlled experiments for several simple network topologies and mobility patterns to test the applicability of STARIMA modeling in wireless environments. We study mobility in a feed forward convergent network, general random mobility in a symmetrical network, and a network with a finite subscriber population. We evaluate the applicability of these models in a controlled environment before applying the models to empirical data.

The rest of this paper is organized as follows: In section 2, we review the STARIMA techniques used in our performance model. In section 3, we present the development of finite-horizon autoregressive models for several simple topologies and mobility patterns. In section 4, we present some numerical results for several call arrival, call holding, and dwell distributions under various mobility patterns. The conclusion is given in section 5.

2 Time Series Models

In this section we review the STARIMA models presented by Pfeifer and Deutsch [16, 17, 19, 20] that are based on common autoregressive models [22].

2.1 STARIMA Process

The space-time ARIMA (STARIMA) class of models presented by Pfeifer and Deutsch [17], and denoted by STARIMA($\boldsymbol{p}_{\lambda_1,\lambda_2,\ldots,\lambda_p}, d, \boldsymbol{q}_{m_1,m_2,\ldots,m_q}$), is defined in vector form as

$$\Delta_d \boldsymbol{z}(t) = \sum_{k=1}^{p}\sum_{l=1}^{\lambda_k} \phi_{kl} \boldsymbol{W}^{(l)} \Delta_d \boldsymbol{z}(t-k) - \sum_{k=1}^{q}\sum_{l=1}^{m_k} \psi_{kl} \boldsymbol{W}^{(l)} \boldsymbol{\epsilon}(t-k) + \boldsymbol{\epsilon}(t). \tag{1}$$

Here, p and q are the autoregressive and moving average orders respectively; λ_k and m_k are the spatial orders of the kth order autoregressive and moving average terms respectively; $\Delta_d \boldsymbol{z}(t)$ is the vector of dth order differences of the observations $\boldsymbol{z}(t)$; ϕ_{kl} and ψ_{kl} are the kth order time lag and lth order spatial lag autoregressive and moving average parameters, respectively; $\boldsymbol{W}^{(l)}$ is the lth order weight matrix; and $\boldsymbol{\epsilon}(t)$ is a vector of random normal errors defined as

$$E[\boldsymbol{\epsilon}(t)\boldsymbol{\epsilon}(t+s)'] = \begin{cases} \sigma^2 \boldsymbol{I}_N & s = 0, \\ 0 & \text{otherwise}. \end{cases} \tag{2}$$

3 Model Development

The evaluation of these models as being appropriate for mobile wireless networks will be based on the method suggested in [17], also known as the Box Jenkins Method [22] which we outline below.

Identification Stage: The first step is to determine whether the space-time series of observations obtained from empirical measurements are stationary. This can be determined by computing the space-time autocorrelation function (ACF) and partial space-time autocorrelation function (PACF) or by using Kalman filtering or other techniques on the data. If the process is not stationary, then the backshift operator can be used to obtain the second-order differences (differences of the differences). Then the ACF and PACF are recalculated and the backshift operator is applied until they display stationarity. Depending on the values of the ACF and the PACF derived from equations (5)and (6), a candidate model is chosen. This can either be a STAR, STMA, or STARMA or STARIMA model.

Estimation of Parameter Stage: After a candidate STARIMA model is chosen, the parameters ϕ and ψ are fitted using MLE estimates using equations (7)-(10).

Diagnostic Checking Stage: The model will be checked against the actual data and simulation models to see that it accurately represents the dynamics of the observed network behavior. Also, if the model is unduly complex or if the model does not adequately fit the planning horizon, adjustments may need to be made. Additionally, we check the autocorrelations and the partial autocorrelations of the residuals to make sure that they are not significant.

4 Numerical Results and Discussion

In this section we study a few common network topologies and simple mobility patterns to evaluate the ability of the space-time autoregressive models to model the spatial and temporal correlations in the system, and assess how good the prediction is based on the model we built. The mobility patterns we study are basic mobility models commonly found in the literature. Analysis of empirical measurements taken for automobile traffic [21] show that higher order weight matrices are not usually necessary in constructing accurate models, so we restrict our studies to first- and second-order neighbors.

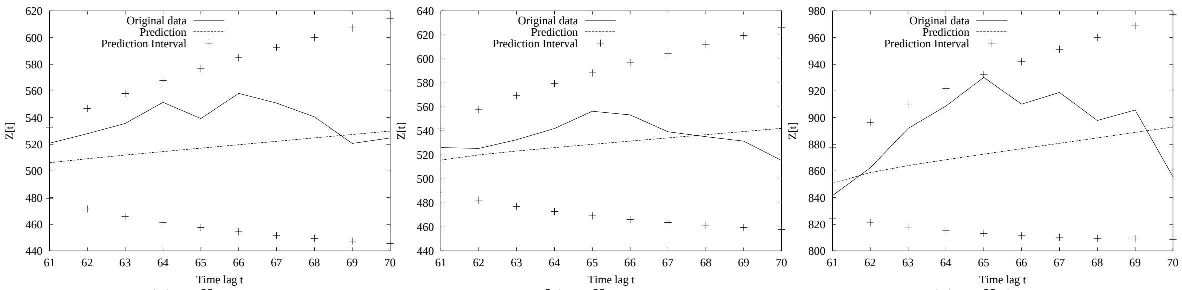
4.1 Experimental Setup

Synthesized traffic traces are generated that represent non-stationary system behavior over a finite time horizon. We model call arrivals as steadily increasing, which can be expressed by a univariate time series: $X_t = 1.8X_{t-1} - 0.8X_{t-2} + \epsilon_t$, where X_0 is the starting number of calls, and ϵ_t is normal (0,1). Subscribers within all cells move at the same constant velocity, with a mean dwell time in a cell which varies in the different scenarios from 15 to 60 seconds. The dwell time for new calls in a cell will be uniformly distributed. Call holding times are Pareto distributed with a mean value of 100s to account for heavy-tailed effects from subscribers sending and receiving data. The number of call requests, both new and handoff, for each cell is sampled every 30 seconds

4.2 Convergent Network

Fig. 1, depicts a small 3 cell network where cell 1 and cell 2 border with cell 3. All three cells have the same traffic pattern, but calls in cell 1 and cell 2 can be handoff to cell 3. The starting point for the call arrival process is X_0 = 2, and the dwell time is 15s. The original data trace is shown in Fig. 2. Prediction coefficients are generated using 60 samples. The zeroth- and first-order weight matrices are

$$\boldsymbol{W}^{(0)} = \boldsymbol{I}, \text{ and } \boldsymbol{W}^{(1)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.5 & 0.5 & 0 \end{bmatrix}.$$



(a) cell 1 (b) cell 2 (c) cell 3

Fig. 5. 10 step prediction

where $\boldsymbol{Y}[i]$ = $\boldsymbol{Z}[i]$-$\boldsymbol{Z}[i-1]$, $\boldsymbol{Z}[i]$ and $\boldsymbol{\epsilon}[i]$ are vectors of observations and errors.

Using equations (7)-(10) to get the MLE estimates, the values for the autoregressive parameters are ϕ_{01} = 0.3627 , ϕ_{11} = 0.6397, ϕ_{12} = 0.3993, and ϕ_{13} = -0.2598.

With the parameters estimated, we check the residual error's spatial autocorrelation and partial autocorrelation. All values measured are smaller than the 95% confidence interval threshold, showing that the model is a good fit.

We use the model to generate a 10 step (300s) prediction as shown in Fig. 5. The dotted line is the point prediction data, the solid line is for the original data, and the points are the 95% prediction interval bounds computed as $\hat{Z}_{60}[i] \pm 1.96\sqrt{i\sigma^2}$, (see [23].)

From Fig. 5 we can make a number of conclusions: (1) The prediction reflects the basic trend of the data, (2) The multi-step prediction curve is smoother than the real data, thus predicting general trends, (3) All the real data falls within the prediction intervals, by which we can conclude that the prediction is good, (4) The prediction error increases with an increase in prediction steps.

4.3 Symmetrical Network Case

For the second scenario, we study a 9 cell symmetrical network as depicted in Fig. 6. User mobility is symmetrical with equal probability of visiting adjacent cells.

Call arrivals and call holding times are the same as for the 3 cell scenario. The mean dwell time is 60s and the samples times are averages, whereas snapshots were taken in the first scenario. The simulation trace is shown for two neighboring cells numbered 1 and 2 (Fig. 7).

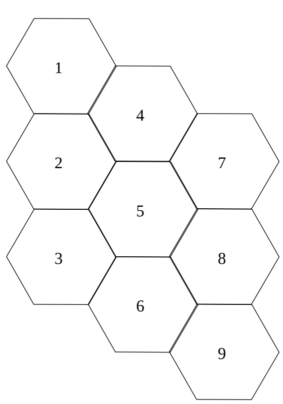


Fig. 6. Symmetrical network

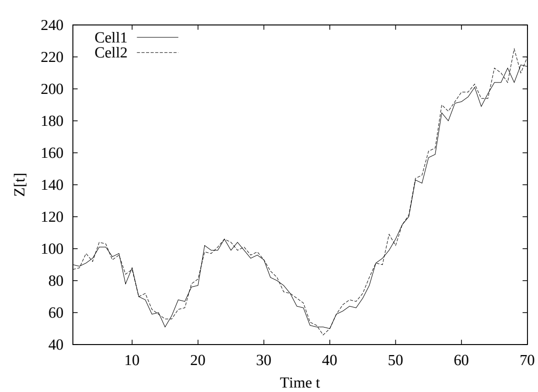


Fig. 7. Symmetrical network data for cell 1 and 2