Automatic Tuning of the Step size in WLMS Algorithms: Applications to EDGE

Lars Lindbom[1], Jonas Rutström*, Anders Ahlén*, and Mikael Sternad*

*Signals and Systems, Uppsala University, PO Box 528, SE-75120, Uppsala, Sweden.
[1]Ericsson AB, PO Box 1038, SE-65115 Karlstad
Lars.Lindbom@ks.ericsson.se, jr@signal.uu.se, aa@signal.uu.se, ms@signal.uu.se

Abstract - **Adaptation laws that estimate time-varying communication channels can be tuned for efficient tracking, by adjusting filters and gains within their structure. However, this requires knowledge of statistical properties that may be poorly known or time-varying. A possible approach is then to adjust the gains of the adaptation law adaptively. The Wiener LMS adaptation algorithm (WLMS) attains efficient tracking by incorporating filters that are attuned to the fading statistics. It includes a gain parameter that will here be adjusted adaptively on line. A WLMS channel predictor with adaptively adjusted gain is then evaluated as a component of receivers for EDGE systems, which use delayed decision feedback sequence estimation. The channel variations for such cases have a wide variety of properties that are tracked by gain adaptation.**

I. Introduction

When adaptively estimating time-varying systems, such as fading mobile radio channels, an appropriate adjustment of the adaptation step size will be crucial. It is desirable to adjust it to optimize the tracking performance, while stability is preserved.

If the parameters to be estimated have time-invariant and known dynamics, the step-size selection can be accomplished by off-line Monte Carlo simulations. However, if the channel parameter properties are unknown or time-varying, then there are three possible approaches:

- *Robust design.* The adaptation law may be designed for a set of possible situations, and its performance is optimized on average over this set, or for the worst case [1, 2]. A robust design may be unsatisfactory if one single estimator is to be adjusted to a very wide variety of situations.
- *Gain scheduling.* A set of adaptation laws can be pre-adjusted for different environments, such as fading rates and SNR's. These properties are measured on-line, so that the appropriate algorithm can be utilized. This approach becomes cumbersome if a high-dimensional space of parameters has to be covered by appropriately designed estimators.
- *On-line tuning.* This can be performed by *indirect adaptation* involving the updating of models of the environment followed by an on-line redesign of the adaptation algorithm. An alternative is to use *direct adaptation* of gains[1].

This paper focuses on direct adaptation.

In the literature, adaptation of the step size has been proposed and successfully applied for LMS algorithms, see e.g. [3],[4]. However, in applications such as e.g. communication over fading channels, the LMS algorithm is frequently inappropriate. A major reason is that the LMS algorithm is most suitable for tracking parameters varying as random walks [5], a statistics that does not describe fading channels well.

It is therefore motivated to investigate other, more flexible, tracking algorithms such as Wiener LMS (WLMS). The tracking MSE can for such algorithms be minimized for a given parameter statistics. Compared to LMS, we may then cope with a larger class of time variations, while maintaining almost LMS complexity. See [5, 2, 6] for details.

In this paper we introduce the variable step-size WLMS (VSWLMS) and investigate the properties of the WLMS algorithm with an adaptively adjusted step-size. As illustrations we consider equalization of EDGE [7] channels and frequency offset compensation. We also compare the adaptively optimized step-size with the optimal step-size that may be calculated analytically my means of recent results in [8].

[1]Combinations of these two approaches, or combinations with robust design and gain scheduling, may also be considered.

II. The WLMS Algorithm

The WLMS algorithm can be summarized as follows. For measurements described by linear regression models

$$y_t = \varphi_t^* h_t + v_t \quad , \tag{1}$$

with h_t being the parameters, φ_t^* the regressors with known correlation matrix $\mathbf{R}$, and v_t the noise, the WLMS adaptation algorithm can be expressed as

$$\begin{aligned} \varepsilon_t &= y_t - \varphi_t^* \hat{h}_{t|t-1} & (2)\\ \hat{h}_{t|t} &= \hat{h}_{t|t-1} + \mu \mathbf{R}^{-1} \varphi_t \varepsilon_t & (3)\\ \hat{h}_{t+k|t} &= \boldsymbol{\mathcal{P}}_k(q^{-1}) \hat{h}_{t|t} \quad . & (4) \end{aligned}$$

Here, $\boldsymbol{\mathcal{P}}_k(q^{-1})$ is constrained to be a diagonal rational matrix with equal stable and causal transfer functions along the diagonal,

$$\boldsymbol{\mathcal{P}}_k(q^{-1}) = \frac{Q_k(q^{-1})}{Q_0(q^{-1})} \mathbf{I} \quad . \tag{5}$$

Above, $\hat{h}_{t+k|t}$ is an estimate of h_{t+k} at sample time t, which may involve prediction ($k > 0$), filtering ($k = 0$) or fixed-lag smoothing ($k < 0$). The superscript $*$ denotes conjugate transpose and q^{-1} is the backward shift operator ($q^{-1} x_t = x_{t-1}$). The algorithm can be regarded as a generalization of LMS that uses orthogonalization with respect to the regressors ($\mathbf{R}^{-1}$ in (3), $\mathbf{R} = \mathrm{E}\, \varphi_t \varphi_t^*$) and filtering of the estimate.

The real-valued gain or step-size μ and the "coefficient filter" $\boldsymbol{\mathcal{P}}_k(q^{-1})$ are design variables. They can be adjusted to minimize the MSE of the tracking error

$$\mathrm{E}\, |\tilde{h}_{t+k}|^2 \ \stackrel{\Delta}{=}\ \mathrm{E}\, |h_{t+k} - \hat{h}_{t+k|t}|^2$$

if a model describing h_t is assumed known. The problem of obtaining an MMSE adjustment of μ and $\boldsymbol{\mathcal{P}}_k(q^{-1})$ is in [5, 6] converted into a Wiener problem by a transformation of signals. The Wiener estimator is obtained by solving a spectral factorization and a linear Diophantine equation iteratively.

For second order AR (possibly integrated) models with real-valued coefficients,

$$h_t = \frac{C(q^{-1})}{D(q^{-1})} \mathbf{I}\ e_t = \frac{1}{1 + d_1 q^{-1} + d_2 q^{-2}} \mathbf{I}\ e_t \tag{6}$$

the optimal polynomials $Q_k(q^{-1})$ are given by the closed form expression

$$Q_k(q^{-1}) = \mu (1\ q^{-1}) \begin{pmatrix} -d_1 & 1 \\ -d_2 & 0 \end{pmatrix}^k \begin{pmatrix} 1 \\ p \end{pmatrix} \tag{7}$$

for $k \geq 0$, where the real-valued scalar p is defined as

$$p = \frac{d_1 d_2 (1 - \mu)}{1 + d_2 (1 - \mu)} \quad . \tag{8}$$

This *Simplified Wiener LMS algorithm*, has only three adjustable parameters, d_1 and d_2 and μ. The parameters d_1 and d_2 should be adjusted to match the actual time variations. In a fading radio environment, one may use a pre-specified robust tuning or on-line adjustment of (6) to the fading statistics (an example of indirect adaptation).

The step-size parameter μ is more difficult to determine. Its optimized tuning depends on a parameter-drift-to-noise ratio which indirectly depends on the properties of the adaptation law itself. This is why an off-line adjustment of WLMS algorithms has to be performed iteratively [5, 6], in general.

It is therefore of interest to adjust μ adaptively on line. We will investigate this issue for situations where $\mathbf{R} = \sigma_z^2 \mathbf{I}$ (white regressors), or when φ_t^* is a scalar (with a not necessarily white regressor).

III. Variable Step-size WLMS

For $\mathbf{R} = \sigma_z^2 \mathbf{I}$, a normalized step-size

$$\bar{\mu} \ \stackrel{\Delta}{=}\ \mu / \sigma_z^2$$

will enter equation (3) in the same way in which the step size parameter affects the LMS algorithm. To adjust μ adaptively we may therefore apply a gain adjustment algorithm that has been developed for use in LMS. Assuming scalar y_t, we here use the variable step-size scheme "c-VSLMS-II-M' considered in [4], which can be expressed as

$$\begin{aligned} \bar{\mu}_{t+1} &= \bar{\mu}_t \left(1 + \rho\, \Re\{\psi_t \varphi_t \varepsilon_t\}\right) & (9)\\ \psi_{t+1} &= \psi_t (\mathbf{I} - \bar{\mu}_t \varphi_t \varphi_t^*) + \varphi_t^* \varepsilon_t^* \quad . & (10) \end{aligned}$$

Above, $\Re\{x\}$ denotes the real part of x, $\bar{\mu}_t$ represents an estimate of the normalized step-size μ/σ_z^2 at sample time t and ρ is a design parameter, typically much less than one. The elements of the row vector ψ_t can initially be set to zero.

By combining the variable step-size scheme (9),(10) with the simplified WLMS structure, in which $\boldsymbol{\mathcal{P}}_k(q^{-1})$ is specified by (5),(7), the variable step-size WLMS (VSWLMS) algorithm for linear regression models with white or scalar regressors is given by

$$\begin{aligned} \varepsilon_t &= y_t - \varphi_t^* \hat{h}_{t|t-1} & \\ \hat{h}_{t|t} &= \hat{h}_{t|t-1} + \bar{\mu}_t \varphi_t \varepsilon_t & (11)\\ p_t &= \frac{d_1 d_2 (1 - \bar{\mu}_t \sigma_z^2)}{1 + d_2 (1 - \bar{\mu}_t \sigma_z^2)} & (12)\\ \hat{h}_{t+1|t} &= -p_t\, \hat{h}_{t|t-1} + (p_t - d_1) \hat{h}_{t|t} - d_2 \hat{h}_{t-1|t-1} & \\ \hat{h}_{t+k|t} &= -p_t\, \hat{h}_{t+k-1|t-1} + Q_k(q^{-1}) \hat{h}_{t|t}\ , \quad k > 1\ . & \end{aligned}$$

The variable step-size scheme (9) is one of several possible schemes discussed in [4] that can be combined with the simplified WLMS.

The variable step-size LMS (VSLMS) is a special case of the one-step ahead VSWLMS predictor, obtained by assuming random walk parameters (i.e. $d_1 = -1, d_2 = 0$, giving $p_t = 0$).

Fig. 1 illustrates the convergence of the step-size μ when tracking a scalar sinusoid

$$h_t = e^{j\omega_o t} \quad , \tag{13}$$

with the VSWLMS scheme. The design parameter ρ was set to 0.08 to balance convergence speed against noise sensitivity. The scalar regressor φ_t^* is here white and circular complex Gaussian, with unit variance ($\sigma_z^2 = 1$). Also displayed is the optimal μ (dashed) obtained by minimizing the tracking MSE presented in Theorem 1 of [8]. We note a good agreement between theory and simulations.

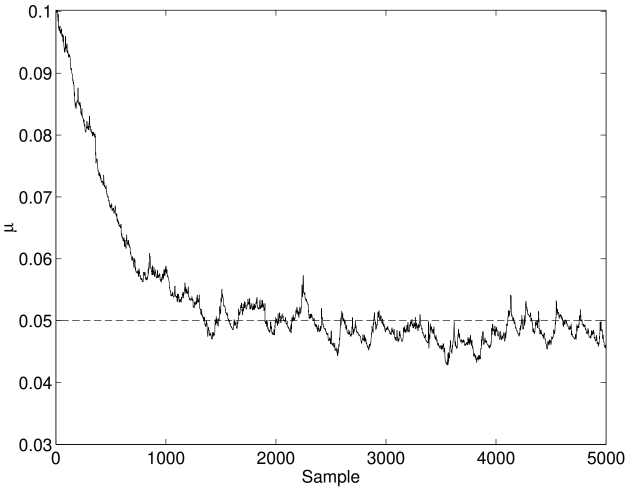

Figure 1: Tracking of the sinusoid (13), $\omega_o = 0.005$, with the VSWLMS algorithm. The SNR is 15 dB whereas $C(q^{-1}) = 1$ and $D(q^{-1}) = 1 - 2q^{-1} + q^{-2}$ in the model (6). The adaptively tuned step size μ (solid) converges to the optimal μ (dashed) after approximately 1000 steps.

IV. Application to EDGE

The block error rate (BLER) performance of a delayed decision-feedback sequence estimator (DDFSE) [9] in conjunction with the VSWLMS algorithm will now be evaluated on multipath fading channels associated with the EDGE air interface. This radio interface is based on GSM, which implies that the symbol rate and the slot format (in terms of symbols) are the same, as illustrated in Fig. 2. In EDGE, the bit rate is adapted to the channel conditions by selecting one of nine different modulation coding schemes (MCS). For MCS five and higher, the modulation is 8PSK with linearized GMSK pulse shaping.

← 156.25 symbols (0.577ms) →

3 tail	58 data1	26 sync	58 data2	3 tail	8.25 guard

Figure 2: The Edge slot format.

An adequate single receiver antenna transmission model for short EDGE/GSM channels, such as Rural Area (RA) propagations [10], can be expressed as

$$y_t = z_t^* h_t + v_t, \tag{14}$$

where v_t is a measurement disturbance, h_t is a scalar time varying complex-valued gain and z_t^* represents a scalar regressor, determined as

$$z_t^* = \sum_{k=0}^{M} g_k s_{t-k} \quad . \tag{15}$$

Above, g_k and s_t denote time-invariant channel taps and transmitted symbols, respectively. For short EDGE channels, the intersymbol interference modeled by (15) is mainly caused by the pulse shaping (partial response) and the receiver (RX) filtering, whereas h_t models flat fading as well as impacts of frequency offsets. In the case of a pure frequency offset, ω_o, originating from a deviation of the carrier frequency from the receiver oscillator frequency, the gain h_t is given by (13). For RA channels, the spectrum of h_t is approximately described by a Rice model [10]

$$\Phi_h(\omega) = A\delta(\omega - 0.7\omega_D) + B\Phi_J(\omega, \omega_D) \quad , \tag{16}$$

for some constants A and B, where $\Phi_J(\omega, \omega_D)$ represents the spectrum of Jakes model [11] with a Doppler frequency ω_D.

The taps g_k are in this presentation estimated by Least Squares (LS) over the sync data interval (Fig. 2), under the assumption that $h_t = 1$, and are then held constant over the entire slot. The estimated channel taps relate to the slot synchronization position that yield the smallest cumulative squared LS residuals. The variance σ_z^2 of z_t^*, required in (12), is also estimated over the sync interval. Tracking of the time varying tap h_t begins in the middle of the sync interval. Decision-directed mode, where known sync symbols are replaced by detected symbols $\bar{s}_t$, is used within the data blocks. The data2 block in Fig. 2 is first detected in the forward direction, while the data1 block is then detected in the backward direction (time-reversal detection). Hence, the equalization will be performed on two half slots.

The considered detector structure is depicted in Fig. 3. The tracker has to perform d-step prediction, since it works on d-step delayed data. The prefilter is designed, for each slot half, as a feedforward filter of an MMSE decision-feedback equalizer (DFE).[2]

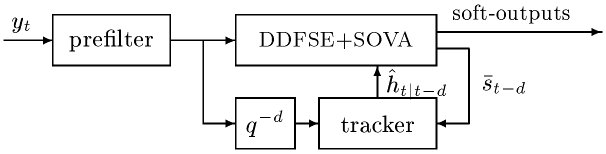

Figure 3: Soft-output DDFSE with d-step prediction tracking of a flat fading channel. The tracker operates on pre-filtered data and delayed symbol decisions $\bar{s}_{t-d}$. Soft sample values for convolutional decoding are here delivered by the soft-output Viterbi algorithm (SOVA) [12].

In the simulations below, a square root raised cosine RX filter with single-side bandwidth of 150 kHz and roll-off factor of 0.15 is used.

Fig. 4 shows the block error rate (BLER) when transmitting over a rather severe fading channel at the 1800MHz band. In this example, the prefilter consist of 15 taps, the DDFSE consist of two MLSE taps and two feedback taps, $M = 3$ in (15), and a decision delay d of three symbols is used. Ideal frequency hopping is assumed.

Therefore, a d-step ahead prediction VSWLMS is used. It is designed for $d_1 = -2, d_2 = 1$ in (6) i.e an *integrated* random walk model. Its performance is compared to that of the variable step size LMS (VSLMS), obtained with $d_1 = -1, d_2 = 0$ (random walk model). For both the VSWLMS and the VSLMS, the design parameter ρ in (9) is set to 0.08, and the variable step-size $\bar{\mu}_t$ is adapted over the slots.

Trackers without adaptive step-size, here denoted WLMS and NLMS, were also evaluated. The normalized LMS (NLMS) algorithm with fixed gain here corresponds to a WLMS with $d_1 = -1$ and $d_2 = 0$ (μ is normalized by σ_z^2). The fixed step-sizes were optimized for zero frequency offset and a mobile speed of 100 km/h, which provided good performance on average over vehicle speeds between zero and 200 km/h. These values were also used as initial values for $\mu = \bar{\mu}\sigma_z^2$ in the algorithms with adaptive step-size.

The VSWLMS algorithm overall provides superior performance. At 200 km/h and 10% BLER in Fig. 4, we obtain an improvement of about 4 dB in comparison with the VSLMS algorithm. Note also that the WLMS with incorrectly tuned (fixed) step-size will provide better performance than the VSLMS algorithm.

Fig. 5 shows the BLER performance when transmitting over a static channel [10], without frequency offset. In this case, tracking is not necessary. It is evident that the WLMS and the NLMS algorithms, both with constant step-sizes optimized for 100 km/h, show a performance loss of about 0.5 dB, or more, compared to the tracking schemes with variable step-size, which have the ability to reduce their step size μ in this case. The performance of the VSLMS and the VSWLMS algorithms are in this case identical, and marginally better than for a receiver without tracking.

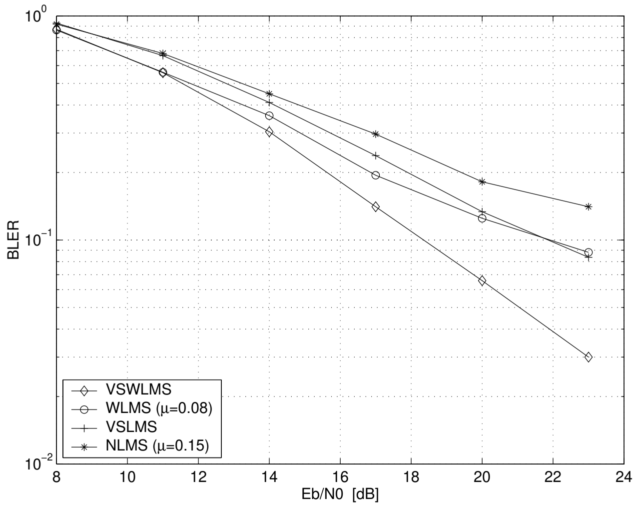

Figure 4: Block error rate for MCS-7 and an RA channel at 1800MHz, at a mobile speed of 200 km/h and a frequency offset of 200 Hz ($\omega_o \approx 0.005$). Comparison between Variable Step-size Wiener LMS (VSWLSM), WLMS with fixed step size, Variable Step-size LMS (VSLMS) and (normalized) LMS with fixed step size (NLMS).

Fig. 6 illustrates the convergence speed of the step size parameter μ at different signal to noise ratios, for $\rho = 0.08$ in (9). Convergence occurs over relatively few slots, which indicates that the adaptation should be able to track variations in the fading environment, the noise level and the speed of the mobile. This issue does however require further investigation.

V. Concluding Remarks

A method for step size adaptation, previously known for LMS adaptation laws, has here been applied to more versatile (simplified) Wiener LMS trackers.

The variable step size WLMS algorithm presented in Section III is limited to tracking parameters in

[2]This receiver structure is similar to the one used in [13] (single branch), but where we have included one-tap tracking and use another soft-output scheme.

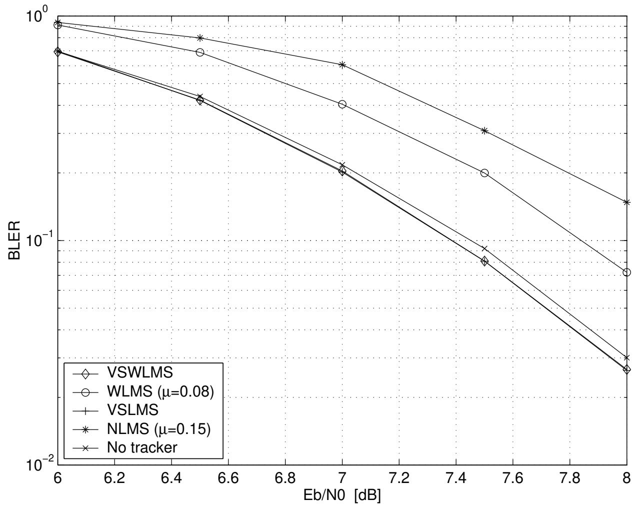

Figure 5: Block error rate for MCS-7 (code rate 0.78 and 8PSK) and a static channel.

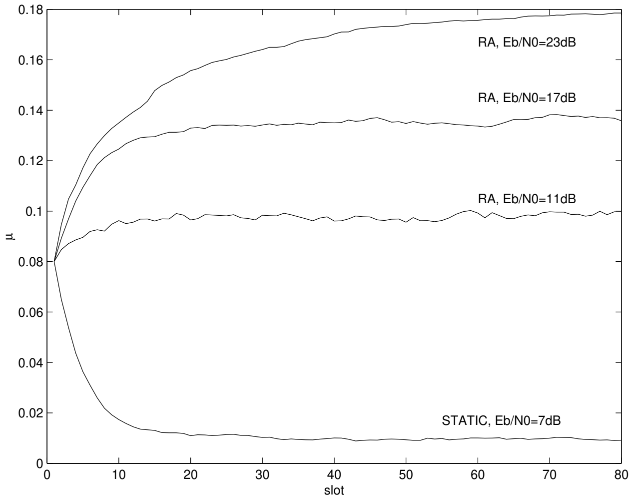

Figure 6: The ensemble mean of 50 realizations of the adaptively tuned step-size μ versus the slot number.

linear regression models with scalar or white regressors. A generalization to colored regressors which can also use adaptively updated ARMA fading models of higher order than two, is under investigation. A generalization can furthermore be conceived to adaptive adjustment of the more general constant gain algorithms recently presented in [6]. Convergence properties are here an interesting topic.

Adaptively adjusted adaptation laws should form a useful addition to the tools of robust and gain scheduling design. Which tool to use in a particular application then becomes a matter of practical tradeoffs between complexity and performance.

References

[1] M. Sternad, L. Lindbom and A. Ahlén, "Robust Wiener design of adaptation laws with constant gains," *IFAC Workshop on Adaptation and Learning in Control and Signal Proc.*, Como, Italy, Aug. 2001.

[2] L. Lindbom, A. Ahlén, M. Sternad and M. Falkenström, "Tracking of time-varying mobile radio channels. Part II: A case study." *IEEE Trans. on Commun.,* vol. 50, pp. 156-167, Jan. 2002.

[3] S. Haykin, *Adaptive Filter Theory.* 3d. ed., Prentice Hall, Englewood Cliffs, NJ, 1996.

[4] Wee-Peng Ang and B Farhang-Boroujeny, "A new class of gradient adaptive step-size LMS algorithms" *IEEE Trans. on Signal Processing*, vol 49, pp. 805-809, April 2001.

[5] L.Lindbom, M.Sternad and A.Ahlén, "Tracking of time-varying mobile radio channels. Part I: The Wiener LMS algorithm," *IEEE Trans. on Commun.*, vol. 49, pp. 2207-2217, Dec. 2001.

[6] M. Sternad, L. Lindbom and A. Ahlén, "Wiener design of adaptation algorithms with time-invariant gains," *IEEE Trans. on Signal Proc.,* vol. 50, August 2002.

[7] A. Furuskär, S. Mazur, F. Müller and H. Olofsson, "EDGE: Enhanced data rates for GSM and TDMA/136 evolution". *IEEE Personal Commun.,* vol. 6, pp. 56-66, June 1999.

[8] L. Lindbom, M. Sternad and A. Ahlén, "Adaptation with constant gains: analysis for fast variations," IEEE *ICASSP 2002*, Orlando, Florida, May 2002.

[9] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation". *IEEE Trans. on Commun.,* vol. 37, pp. 428-436, May 1989.

[10] GSM 05.05, "Radio transmission and reception," European Telecom. Standards Institute (ETSI).

[11] W. C. Jakes, "Multipath Interference". In W. C. Jakes ed. *Microwave Mobile Communications.* Wiley, New York, NY, 1974.

[12] J. Hagenauer, "Source-controlled channel decoding". *IEEE Trans. on Commun.,* vol. 43, pp. 2449-2457, Sept. 1995.

[13] S. Ariyavisitakul, J. Winters and N. Sollenberger, "Joint equalization and interference suppression for high data rate wireless systems". *IEEE J. Select Areas Commun.,* vol. 18, pp. 1214-1220, July 2000.