Estimation of Link Quality and Residual Time in Vehicular Ad-Hoc Networks

Nikoletta Sofra and Kin K. Leung
Electrical and Electronic Engineering Dept.
Imperial College
London SW7 2BT, United Kingdom
Email:{nsofra, kin.leung}@imperial.ac.uk

Abstract—**High node mobility and transient connectivity in Vehicular Ad Hoc NETworks have introduced numerous challenges in the design of efficient communication protocols for these networks. Toward the goal of efficient design of such networks, the focus of this work is to develop methods that will allow the estimation of a wireless link's future quality and the remaining (residual) time for which this will remain efficiently active and useful for data transmission. Such information can serve as key input to the routing algorithms and data dissemination for VANETs. Specifically, we propose a cross-layer approach to the link estimations by introducing a method that utilizes the received signal strength of data packets. Small scale fading induced by the relative movement of the nodes, limited knowledge of the signal, caused by the fact that the received energy metric will be numerically available only when packets are sent, and the small number of available samples of the metric make the problem challenging. To overcome these, a signal processing technique, Empirical Mode Decomposition is used along with robust regression for the prediction of link quality and residual time. The validity of the proposed methods are tested and validated via extensive simulations.**

Index Terms—**Vehicular Ad Hoc Networks, link quality, link residual time, mobility prediction, received signal strength**

I. INTRODUCTION

THE field of Vehicular *Ad hoc* NETworks (VANETs) is constantly drawing research attention due to the wide range of technological applications that they have and the special characteristics and technical challenges that they exhibit. They offer a tempting solution for inter-vehicle communications without the need for infrastructure and compared to existing technologies such as cellular GPRS or 3G networks, they reduce the cost by employing the *ad-hoc* wireless paradigm. Constituting a subclass of Mobile *Ad hoc* NETworks (MANETs), VANETs are self configuring networks of mobile routers connected via wireless links, and need to rely on the nodes themselves to perform network management functions. They are however much distinguishable from most generic MANETs, in the sense that their inherent attributes make the design of efficient communication protocols much more challenging.

One of the biggest challenges in VANETs is the high mobility of the nodes, which results in frequent fragmentation and short-lived links. Prior knowledge of existing links' residual time, their future quality or notifications of imminent breakage of a link could assist the routing decision to a great extent. The availability of such information is crucial for the selection of appropriate routes, or the forming of clusters within the network. More importantly, it is essential in seamless handoff processes, where the discovery of a new route can be initiated before the current one breaks to avoid disruption of data flow. The study and development of techniques for predicting such link quality information is the focus of this work.

A method using global positioning system (GPS) information to predict the link expiration time is presented in [1]. A free space propagation model is assumed and the duration of a connection between two nodes is then calculated based on the transmission range, speed, direction and position of the nodes. However, being within range does not necessarily guarantee radio connectivity and closer physical distance does not always imply better link quality, especially in an urban environment.

The use of link quality feedback in terms of signal strength is considered in [2], where the metrics of signal strength and location stability are introduced. Beacon messages are exchanged between the nodes to maintain connectivity. These are sent at regular intervals, once every time quantum. Signal strength is taken into account by calculating the exponentially smoothed average of the signal strength of every packet received by a specific sender. In order to monitor the temporal stability of the link they record how long beacons have been continuously received with a strong signal. Similarly, [3] use the smoothed SNR value to evaluate the link. Nevertheless, no details as to how or how often the signal-to-noise ratio

(SNR) is sampled, or results on the success of these monitoring techniques are given in the afore-mentioned studies. These approaches cannot provide any information other than the exponentially smoothed average value of SNR from previous samples. This cannot be utilized to predict the link residual time, which is defined as the remaining time interval until the link quality (e.g., in terms of received signal power) drops below a given level. Moreover, exponential smoothing gives a smoothed but lagged version of the available data, rather than an extrapolation into the future.

The novelty of this work lies in utilizing information available in the physical layer for estimating the future link quality and the residual time of the link between a given pair of nodes that lie in the proximity of each other using signal processing methods when sparse samples of the received energy are available. The proposed techniques are validated by running computer simulations for short duration links (approximately 10 seconds to 2 minutes) that are affected by small scale fading. More precisely, the signal strength of all packets transmitted by a node is monitored by each neighboring node, so that a time series of received energy samples is formed. The time series is then used by each node individually, producing estimates concerning the link between the two nodes. Every time a new packet is received, the method is applied so that new estimates can be acquired. Empirical Mode Decomposition and robust regression techniques are used for the estimation.

The rest of the paper is organized as follows. We firstly address in Section II issues concerning realistic simulation of the temporal changes in the channel and sampling. Sections III and IV present the new techniques based on regression and Empirical Mode Decomposition. Results revealing the performance of the proposed techniques are presented in Section V and finally we conclude.

II. SIMULATION PRODUCED DATA AND SAMPLING

In order to evaluate any prediction method, we need to produce a series of data that resemble the temporal changes of the actual received signal strength as accurately as possible. It is very important that the small scale fading (high frequency noise), caused by the relative movement between transmitting and receiving nodes as well as the temporal changes in the channel, is depicted in our test data. At this point it must be clarified that it is the large scale phenomenon that we are trying to capture and predict, namely the change in the average value (either attenuation or amplification) caused by change in the relative distance between sender and receiver or by loss/disturbance of line of sight or other changes in the environment. However, changes in the signal due to small scale fading are not to be neglected as they will greatly affect the performance of the algorithm. The reason for this is that small scale fading may cause rapid fluctuation of the signal by as much as three or four order of magnitude (30-40 dB) when the receiver is moved by only a fraction of a wavelength [4], introducing

considerable noise. Therefore, we employ the Rayleigh fading channel model, and specifically, the Jakes channel model for the Doppler spectrum.

In order to model the distance dependence of the received energy, we attenuate the signal by assuming that the received strength is declining proportional to d^a where a is the attenuation exponent factor and d is the distance between the sender and receiver. Details about the modeling of the distance are described in Section V.

As far as sampling is concerned, we distinguish two different situations. The first is when there already exists a data flow on the link so the monitoring can actually be very frequent. As for the other case, we have to rely on "hello" messages, sent by the nodes to retain connectivity or on overhearing packets transmitted by the neighboring nodes on other links via the wireless medium. The regularity and duration of the packets received will greatly affect the performance of any algorithm used, since these constitute the time series that will act as input for the method. More specifically, when the packet is longer, our data will be more insensitive to the noise induced by small scale fading, as this will be averaged out to a certain extent. The data will also be less noisy for higher speeds due to the larger value of Doppler frequency. In addition, the data will be aliased, since the frequency with which packets are sent, i.e., our sampling frequency, does not satisfy the Nyquist criterion for the frequency of the noise. The initiation of the method occurs after packets have been received for 3 seconds. Short lived links and infrequent reception of packets may mean that in certain cases the maximum number of samples available as inputs to the prediction method is 15 or lower, while this number increases to hundreds if the relative speed between the vehicles is low and if packets are transmitted more frequently.

III. EMPIRICAL MODE DECOMPOSITION

Because of the noise that heavily influences our data, we apply a denoising (i.e. noise removal) process. This involves smoothing of the available data by using Empirical Mode Decomposition (EMD) [5]. Using EMD for the signals in question is well justified, since this method considers oscillations at a very local level. It is in this way suitable for non-stationary signals, such as the received energy for mobile nodes, as the speed with which the channel changes, and hence the Doppler frequency is variable. Furthermore, it also applies for non-periodically distributed samples in time, so missing packets or packets that are not sent periodically can be handled by the method.

Very briefly, EMD is a signal processing technique that comprises of successive steps gradually dissecting the signal into *local detail d(t)* (high local frequency part) and *local trend m(t)* (low local frequency part). The procedure, presented in [5] and summarized in [6], includes the computation of the lower and the upper envelope of the signal, $e_{min}(t)$ and $e_{max}(t)$ respectively, by interpolating between local *extrema*. The local trend can then be identified as the mean of the envelopes:

$$m(t) = \frac{e_{min} + e_{max}}{2}, \qquad (1)$$

while the local detail is what remains of the signal:

$$d(t) = x(t) - m(t). \qquad (2)$$

A *sifting* process is then applied by iterating the trend extraction on the local detail until certain conditions are met. The resulting local detail after the sifting process is the first *intrinsic mode function*, c_1. The intrinsic node function is then subtracted from the signal and the whole procedure is reapplied, so that in the end the signal is dissected into a number of intrinsic functions and the *residual*, which is the global mean linear trend of the signal:

$$x(t) = \sum_{i=1}^{N} c_i + r_n \qquad (3)$$

The smoothing process we use includes decomposing the available samples with EMD, using the criteria described in [6] for the halting of the sifting process, up to the point where oscillations occurring because of small scale fading are eliminated, to the possible extent. In practice, this is achieved by adding the last extracted intrinsic mode function with the residual, as given being the output of the EMD algorithm. The overall effect is that of *locally* smoothing the data.

Fig. 1 shows an example on how EMD smoothing will denoise our data: the channel state at all times (considered unknown) is depicted, together with the samples available to us. The EMD smoothed version of the samples is shown together with the exponentially smoothed version for comparison (a typical value for the smoothing parameter, α=0.15 has been chosen).

IV. REGRESSION TECHNIQUES

A. Critical Points Identification

In order to apply the regression model on the samples, it is of vital importance that turning points on the data are promptly identified. Turning points are defined as the *extrema* of the received energy on the large scale, implying change of direction of the relative movement between the two nodes. We distinguish two phases, the *improving* phase when nodes move closer to each other and as a result the link quality improves, and the *deteriorating* phase, when nodes move apart from each other. Identification of such points is crucial, as any data before the turning point need to be discarded before applying the regression algorithm. Correct identification should occur as soon as possible. The process of determining those critical points, is however not trivial, as the noise is contained in the data in a way that the zero derivative method would yield false results. The method used involves denoising of the time series using EMD and identification of the turning points on the denoised signal based on a combination of a simple differentiation method with localization of extrema based on their surroundings.
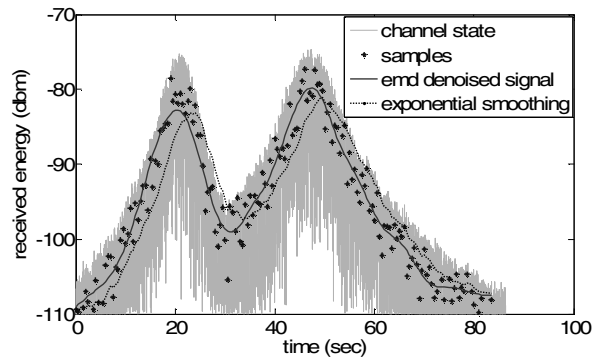

Fig.1 Comparison of EMD and exponential smoothing denoising. The temporal state of the channel and the samples if packets are received periodically every 500 msec are sent are also plotted.

B. Applying Regression

As it is the large-scale path loss aspect that we want to derive out of the sampled data, we can apply regression, treating small scale fading and other random variations in the received signal strength and measurement errors as noise. Regression techniques consist of assuming a functional relationship to exist among the data and then estimating the parameters of this relationship [7].

According to [4], both theoretical and measurement-based models indicate that the received energy in decibel for outdoor channels varies according to the log-distance path loss model, which is an extended generalization of the Friis equation:

$$E_r = E_t(PL_0 + 10n \log_{10}(\frac{d}{d_0})) + n_i, \qquad (4)$$

where PL_0 depends on the frequency, antenna gains and other system loss factors, n is the path-loss exponent depending on the specific propagation environment (e.g. existence of obstructions), d is the separation distance and n_i is the noise caused by small scale fading. Assuming all these factors except for distance as constant, we get

$$E_r = c + a \log_{10}(d) + n_i \qquad (5)$$

If we assume constant speed over certain time intervals, thus having

$$u(t) = u_{const}, t \in T_i, \qquad (6)$$

or equally accept that speed will not dramatically change, so that we can average over this time interval:

$$u(t) = \overline{u}, t \in T_i, \qquad (7)$$

then we finally have

$$d = ut, t \in T_i. \qquad (8)$$

Putting (8) into (5) leads to modeling the received energy according to the following equation:

$$E = a \log_{10} t + b + \varepsilon, \qquad (9)$$

where ε accounts for the error introduced by variations of other forms that are not taken into account by the log distance model (including small scale fading, errors of measurement and sampling biases).

Utilizing given pairs of energy values and the time instances on which these were sampled (i.e., measuring the

energy level of the signal in received packets and keeping track of the times on those where received), we can regress on (9), acquiring values for a and b. We can then use these values to obtain estimates of the energy level at some time (t_f) in the near future or the link's residual time until the value of the received energy drops below a specified threshold γ, according to the following equations:

$$\hat{E} = a \log_{10} t_f + b \tag{10}$$

$$L\hat{R}T(t) = 10^{\frac{\gamma-a}{b}} - t \tag{11}$$

We compare two different methods of regression, namely simple regression and robust regression. Simple regression gives an estimate of a and b by solving

$$[a,b] = ((\log_{10}(\underline{T}))^T (\log_{10} \underline{T}))^{-1} (\log_{10}(\underline{T}))^T \underline{S}, \tag{12}$$

where $\underline{T}$ is the vector of times and $\underline{S}$ is the vector of the samples of energy received at these times. Eq. (12) is well known to minimize the sum of squares errors on the observed values:

$$S = \sum_{i=1}^{N} (\hat{E}_i - E_i)^2 = \sum_{i=1}^{N} (a \log_{10} t + b - E_i)^2, \tag{13}$$

where N is the number of available samples.

Robust regression is an iterative form of simple regression, where pairs of independent and dependent variables are assigned weights based on the residuals from the previous iteration. This way, the parameters acquired are less sensitive to outliers that may exist among the sampled data. Specifically, we try to minimize the weighted sum of the square errors:

$$S = \sum_{i=1}^{N} w_i^2 (E_i - \hat{E}_i)^2 = \sum_{i=1}^{N} w_i^2 (a \log_{10} t + b - \hat{E}_i)^2, \tag{14}$$

thus solving

$$[a,b]^{(t)} = ((\log_{10}(\underline{T}))^T \underline{W}^{(t-1)} (\log_{10} \underline{T}))^{-1}$$
$$(\log_{10}(\underline{T}))^T \underline{W}^{(t-1)} \underline{S}, \tag{15}$$

where W is the diagonal matrix containing all weights.

For each iteration the weight at each point is a bisquare function of the residual of the previous iteration:

$$w_i = \begin{cases} 1 - (r_i)^2, |r_i| < 1 \\ 0, |r_i| > 1 \end{cases},$$

$$r_i = \frac{e_i}{tune * s_i * \sqrt{1 - h_i}}, e_i = E_i - \hat{E}_i \tag{16}$$

where $tune$=4.685, s is an estimate of the standard deviation of the error term and h_i is the leverage value from a least squares fit.

V. PERFORMANCE RESULTS OF THE PROPOSED METHODS

A. Settings

The methods are evaluated by considering a pair of nodes that move relative to each other. Packets are assumed to be sent periodically with the intervals ranging from 0.25 to 1

second. These may be either 'hello' packets or data packets being exchanged within the network, whose energy is being monitored by the nodes. Each packet transmission lasts 5% of the time interval, so that there is constant traffic load. This means that packets transmitted more frequently will last less and thus the time series will contain more samples but it will also be noisier, as explained in Section II. The relative speed between nodes ranges between 1 and 8 m/sec (uniform distribution). The value chosen from the uniform distribution is contaminated with additive white Gaussian noise (SNR=10). This reflects a situation where the movement is quite stable in time, but it is realistic, since speed varies for the whole duration of the movement. The frequency used to calculate the Doppler effect is that of IEEE 802.11g, namely 2.4 GHz. The threshold for efficient communication is -110 dBm. Any samples below the threshold are considered not received and not processed by the method. The initial distance between the nodes for sections C and D is uniformly distributed between 1 and 100 m. All simulations were run on Matlab.

B. Critical Points Identification

We hereby evaluate the proposed method of finding turning points. Table I shows turning points that were missed by the algorithm as a percentage of the total turning points for different time intervals between packets. Moreover, the percentage of false alarms (the method indicated a critical point when there was none) over the total alarms is given. False alarms should be avoided, but in general they do not influence the method as a whole that much. We can observe that the method yields more false alarms when more frequent but shorter packets are received. This was expected, as these data contain more noise. Because our samples are sparse in time, we can not expect to exactly match the critical point. The sensitivity is the time interval between samples. Fig. 2 shows the absolute mean error in the detection of the time on which the critical point (t_{cp}) is observed relatively to the time interval (T) used:

$$r = \frac{|t_{cp} - \hat{t}_{cp}|}{T} \tag{17}$$

The distance between sender and receiver is derived through a model that assigns values for the relative speed between the two nodes as described in part A of this section. Nodes are initially placed in random positions within range. The minimum distance that will be achieved between them is uniformly distributed between 1 and 50 m. Changes in the direction of the relative movement are allowed only after a considerable amount of distance has been covered maintaining the previous direction. This model successfully reproduces the changes in distance between vehicles that move in a sequential way, one following the other, coming closer and further to each other as their relative speed changes over time, vehicles moving almost randomly in a tactical environment or simply vehicles passing past each other.

The results presented are the mean values for 400 different time series obtained by simulations.

TABLE I

Time Interval (sec)	0.25	0.5	0.75	1.0
Missed Turning Points (%)	0	0	0	0.5
False Alarms (%)	2.5	0.25	0	0

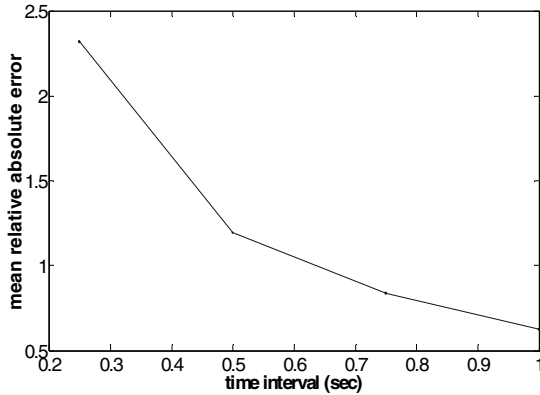Percentage of false alarms and missed turning points using the Critical Points Identification Method.

Fig.2 Accuracy in critical points identification in terms of the mean relative error, as this is defined in (17), plotted versus the time interval between received packets.

C. Prediction of Received Energy

The prediction is made for the time that the next packet is expected to be received, thus after one time interval. Since we aim at predicting the average received energy, the true value at the time of prediction is the average received energy over a time period equal to half of the time interval, so that high frequency noise is averaged out for the calculation of the true value. Prediction at each time point is based on the packets received so far.

The mean absolute error for different values of the interval between received packets is shown on Fig. 3. The methods compared are simple and robust regression on raw and EMD filtered data.

The results presented were obtained by running simulations for 800 different time series. We have averaged over the mean error of the time series that corresponded to same time intervals between packets.

We note that the best performing method for smaller time intervals is using robust regression on EMD filtered samples. The smaller the time interval, the more packets are received, thus enabling better estimation. However, for bigger time intervals, the packets received last longer, as we have made the assumption of constant traffic load. After a certain point, when the duration of the packets becomes large enough, noise tends to be averaged out to a certain extent. This is the reason why robust regression on raw data will perform better than robust regression on emd denoised data as the time interval increases. The denoising process achieved by EMD will rather distort than improve the quality of the data (it should be reminded here that robust regression entails denoising in itself).
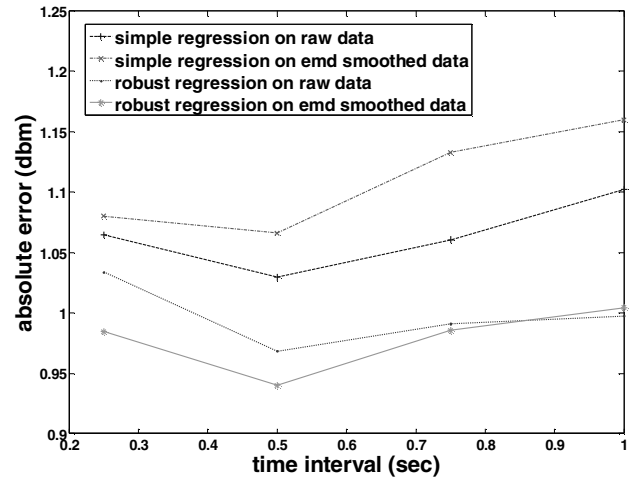
Fig. 3 Mean absolute error of received energy prediction plotted against the time interval of the packets that are periodically received.

D. Residual Link Time (RLT) Prediction

It is generally very intricate to numerically evaluate an RLT prediction algorithm, because there is no specific time point when a receiver will suddenly stop receiving messages. The receiver's efficiency in successfully collecting packets from a specific sender will deteriorate as the channel conditions worsen and as the distance between them increases. However, due to small scale fading we cannot distinguish a certain moment in time from which communication will fail, as deep fades in received energy may cause a packet to be lost although successive packets will be efficiently received. What is more, a receiver's sensitivity is not a strictly set value.

In order to calculate the RLT error, a nominal value of the residual time is necessary. We fit a logarithmic function to the temporal condition of the channel state and assume that link failure occurs on the time point T_f on which the fitted function drops below the specified threshold for successful communication. What is meant by channel state is the received energy in dBm as a function of time, had the transmitter been transmitting continuously. This information is available in our simulations. Fig. 4 further clarifies the definition of T_f. An example of how the received energy varies in time is given. Note that this is not the received energy in the receiver, as our samples are sparse, but a depiction of the channel state. Given this vagueness in defining actual RLT, a certain level of error should be allowed when predicting the link's residual time.

Mean absolute error for RLT estimation based on raw and emd filtered data versus the time interval between received packets is plotted in Fig. 5. Robust regression has been used. Estimation based on EMD filtered data performs better in every case. As the time interval, and therefore the duration of the packets, becomes longer, our data are less noisy, as explained in the previous sections. For this reason, estimation based on raw data tends to approach estimation based on the denoised data for bigger time intervals between packets.

As more packets are received and our time series grows

longer, RLT estimation improves. We have added a convergence criterion to decide whether the RLT estimation made is acceptable or not. In the average case, the algorithm converges after approximately 10 seconds.

The results presented were obtained by running simulations for 1000 different time series. We have averaged over the mean error of the time series that corresponded to same time intervals between packets.
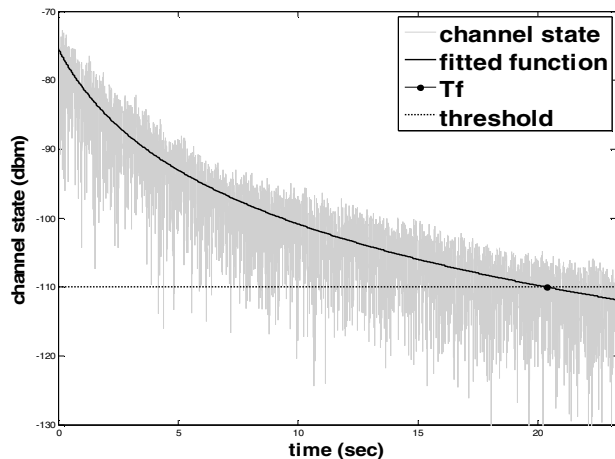

Fig.4 The intricacy of calculating actual RLT. Our error estimations are based on T_f being considered as the link breakage point.
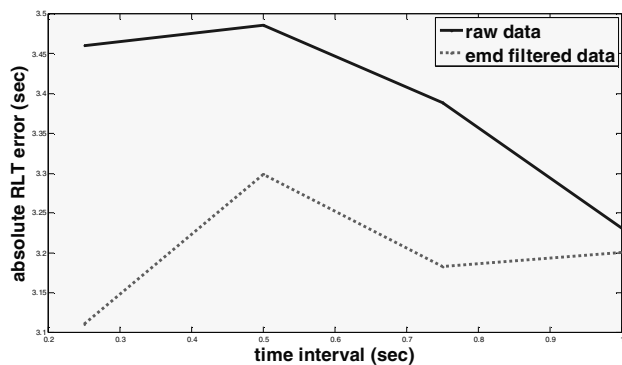

Fig. 5 Absolute RLT prediction error versus time interval between received packets.

VI. UTILIZATION OF THE METHODS FOR THE DESIGN OF NETWORK PROTOCOLS

Of equal importance to the prediction of the RLT of the link or the future level of the received signal strength is the way that these will be utilized to support routing decisions. The algorithms developed need to be capable of indicating the right time point for a handoff and realizing it, as described in [2]. The term handoff in the routing context describes the process of discovering and start using a new route just before the current one expires, so that no gap in communication is observed. Ideally, an interruption on any of the paths comprising the route should pass unobserved at the higher layer, since the routing protocol should have regulated this by switching paths before the interruption occurs. In this way, seamless QoS can be offered at the application level. Additionally, during the construction of a new route, links with higher residual time and/or better

expected link quality should be preferred. A very detailed analysis of adapting existing routing algorithms (both proactive and reactive) for MANETs to this direction is given in [1], if the RLT can be calculated.

Estimation of RLT is also substantial for the design of clustering algorithms, where the prediction of the link residual time can be useful in deciding whether a node will become member of the cluster.

VII. CONCLUSIONS AND FUTURE WORK

In this work we underlined the importance of predicting the future link quality in terms of received signal strength and residual time of wireless links between a pair of moving nodes for the enhancement of routing algorithms for vehicular ad-hoc networks. The method developed was based on sampling the signal strength of the packets received and was tested via simulations depicting realistically the randomness of the channel state. No assumption of knowledge of the speeds and positions of neighboring nodes has been made. The method uses the packets that are transmitted over the network ("hello" messages and packets transmitted on existing data flows) and thus incurs little bandwidth overhead.

Future work includes further development and study of estimation methods, including non-linear parameter estimation for estimating the initial distance between nodes, study of performance for different traffic loads, assumption of non-periodically distributed samples in time, extension of RLT to route residual time, extension of proposed methods to predicting link quality and residual time based on signal-to-interference-plus-noise ratio (SINR) measurements; integration of the prediction methods into efficient routing protocols and data dissemination algorithms for VANETs

REFERENCES

[1] W. Su, S. Lee and M. Gerla, "Mobility prediction and routing in ad hoc wireless networks," *Int. J. Network Management*, vol. 11, no. 1, pp. 3-30, 2001.

[2] Hsin-Mu Tsai; Wisitpongphan, N.; Tonguz, O.K., "Link-quality aware ad hoc on-demand distance vector routing protocol," 1st Int. Symp. on *Wireless Pervasive Computing*, 2006, p.6.

[3] R. Dube, C. D. Rais, W. Kuang-Yeh, and S. K. Tripathi, "Signal stability-based adaptive routing (SSA) for ad hoc mobile networks," *IEEE Personal Commun.* [see also *IEEE Wireless Communications*], vol. 4, no. 1, pp. 36-45, 1997

[4] T.S. Rappaport, *Wireless Communications: Principles and practice*, 2nd Ed., Prentice Hall, 2002

[5] N. E. Huang, et al, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. of Mathematical, Physical and Engineering Sciences* vol. 454, no. 1971, pp. 903-995, 1998.

[6] G.Rilling, P. Flandrin, and P. Gonçalves, "On Empirical Mode Decomposition and its Algorithms," IEEE Signal Processing Letters, vol.11, no 2, p.112, 2004.

[7] S. Makridakis and S.C. Wheelwright, *Forecasting: Methods and Applications*. USA: Wiley/Hamilton, 1978.