www.czasopisma.pan.pl PAN www.journals.pan.pl
POLSKA AKADEMIA NAUK
jet INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2010, VOL. 56, NO. 4, PP. 451–456
Manuscript received May 10, 2010; revised July 2010. DOI: 10.2478/v10177-010-0061-9

Trajectory Estimation for Wireless Mobile Networks Using Polynomial Regression

Mohammad Al-Hattab and Johnson I. Agbinya

***Abstract*—Arbitrary and random motion of mobile ad hoc network nodes while communicating results in frequent topology changes and multiple disconnections of links. This dynamic environment challenges the routing of data from the source to the destination and imposes the need for prediction models to track these changes, and then determine future topology of the network. The prediction of network mobility into the future will reduce the frequency of location updates for geographical routing protocols. Moreover it will reduce route request delay and the frequency of route updates in topology based protocols. This paper proposes a predictive model called polynomial regression trajectory estimation. This model is based on the regular behavior of nodes and uses polynomial regression to allow each mobile node to estimate its future locations as a function of time. The estimated locations will be disseminated to the network so that nodes can use them to estimate the future topologies of the entire network. The efficiency of the proposed model has been evaluated by MATLAB simulation.**

***Keywords*—Topology, prediction, polynomial regression.**

I. INTRODUCTION

MOST mobile nodes show some sort of regularity in their movement [1]. The regularity of the behaviour of nodes can also be seen in many aspects, for example, the volume of traffic generated by a node and the data sent and received shows a form of regularity. This may differ from node to node but within the same node it tends to converge to an average value or to a certain pattern.

The regularity of mobile nodes appears clearly in their movement. For example, a mobile node carried by a vehicle travelling along a street will maintain its current speed and direction for a certain period of time. A plane flying in the sky would maintain regularity in its movement, by flying at a constant speed and toward a given direction for a period of time. This regularity of movement for the node beside the knowledge of its mobility status such as current location, speed, direction of movement and path enables the node to predict its future. Therefore, the future topology of the mobile network could be predicted too.

Recently, several reasons have enhanced the employment of location information in network applications, such as routing decisions, nodes localization and topology estimation. For example, the advance of navigation systems, the low cost of navigation devices and the ease of integrating these devices into most types of mobile node, all contribute towards the deployment of Global positioning system (GPS) [2] receivers wildly in mobile devices and to use them in different applications.

Any GPS enabled node can use navigation software to calculate a possible path to the destination, and determine the coordinates for many points along that path. If we can find a mapping between the coordinates on that path and the time when the node is going to pass these points, we would be able to estimate the future locations of the mobile node as a function of time. Nodes in the network can exchange these estimates to use them in predicting the future topology of the network.

The knowledge of the network's future topology will improve the throughput and the delivery of data in the network. Moreover, it will reduce the control messages and the packet overhead.

Many applications in wireless sensor networks and mobile ad hoc networks require the knowledge of the exact location of nodes [3]. For example, in navigation, a node needs to know its location to calculate a route to the destination using digital maps. The location of nodes can also be used in routing [4]. Routing based on location of nodes called geographic routing. Many algorithms and protocols were proposed to use geographic routing such as DREAM [5], SLURP [6], LAR [7], and GPSR [8].

The node obtains its own location information through GPS capability which is by the date of this paper the only fully functional Global Navigation Satellite System (GNSS) [9]. The system utilizes a constellation of at least 24 medium earth orbit satellites that transmit precise microwave signals. It enables a GPS receiver to determine its location, speed, direction, and current time using mathematical process called trilateration [2], [10]. Nodes may also obtain their location through different localization methods such as in [11] and [12].

In geographical routing, a node must know the location of the destination node to be able to deliver data to that destination. Location information could be stored in the node itself or somewhere in the network in a service called location service. So, each node must inform other nodes or the location service in the network about its location to make geographical routing possible.

The dissemination of location information in the network can be accomplished by two different methods, flooding based method or rendezvous based method [13]. In flooding based method some algorithms propagate the location information periodically to the network such as DREAM [5] and SLURP [6] which disseminates the update when the node moves outside a pre-defined region. In rendezvous method such as in [14] all nodes in the network agree upon a mapping that maps

M. Al-Hattab and J. I. Agbinya are with the Centre for Real Time Information Networks (CRIN), Faculty of Engineering and Information Technology, University of Technology, Sydney (UTS), PO Box 123, Broadway NSW 2007, Australia (e-mails: malhatta@eng.uts.edu.au, Johnson.Agbinya@uts.edu.au).

each node in the network to one or more nodes. The mapped-to nodes are the location server where location updates are stored and retrieved.

In both flooding and rendezvous methods the network will be more bandwidth utilization efficient if the dissemination of location updates is not very frequent. On the other hand less frequent updates will lead to less accurate location information. So, we propose a model that disseminates less periodic updates with a series of estimated locations in each update to achieve both requirements.

II. RELATED WORKS

In previous work [15], [16] we have proposed a localization technique that uses street maps to estimate the location of mobile nodes when they are carried by mobile vehicles. Each node estimates the locations of all other nodes in the network, based on update messages exchanged periodically by all nodes; these updates contain location, speed and direction.

Many other algorithms have been proposed and developed to use location information for mobility prediction and routing. Dead reckoning technique [17] and [18] uses location information to predict node's mobility in mobile ad hoc networks. A mobile node samples its own location periodically and constructs a model of its movement. Two successive location samples taken at known times to calculate the speed in both X and Y direction, then the calculated speed is used to estimate the next location. The distance travelled is calculated and compared to a threshold value. This threshold is to determine the allowable error in the system. The complexity of the model depends on the predictive ability of the mobile node. The node disseminates its current location as well as this model in the network. Every other node in the network uses this information to track the location of this node.

The authors in [1] proposed an algorithm to predict the expiration time of the link between two adjacent nodes. The routes are then reconfigured before the expiration time of the link reached in order to prevent disconnection. The algorithm calculates the expiration time for the link using the speed, the direction and the coordinates for both nodes. This algorithm obtained the information about the location and the speed by piggybacking them during live connections. If the nodes do not communicate very often then the information is not accurate and may lead to inaccurate estimation.

Location Prediction Based Routing (LPBR) [19] proposed to reduce the number of route discovery uses flooding based route request to find a route to the destination. Route Request packet gathers location information for each node traversed, and this information is stored in the destination. Route Reply that contains a path to the destination is traversed to the source node. If the destination node doesn't receive a data packet it constructs a global topology of the network using the recent gathered location information and sends a route to the source to use it in delivering data.

Autoregressive model using Kalman filter was used by several authors to estimate the mobility of the wireless networks [20], [21] and [22]. The Yule-Walker equations [23] and Kalman filter [24] are used to estimate the mobility of wireless network in [21]. A given mobile state at time n can be used to estimate the mobile state at time $n+1$. States n, $n-1$, $n-2\ldots$ are used to calculate the state variables.

In [20]the authors proposed a prediction model called hierarchal location-prediction model (HLP) to predict the future trajectory of the mobile node in ATM cell networks.

Cross correlation and pattern matching are used to predict the future state of the link in wireless networks in [25]. The node monitors and stores a series of signal to noise ratio (SNR) measurements for its neighbours. These measurements are cross-correlated with a certain values of query to predict a future value of the SNR of the link.

A GPS-free node localization in mobile wireless sensor networks [12] is an algorithm proposed to estimate the locations of nodes in wireless networks. This algorithm is GPS-free and nodes have a compass pointing to a common direction. Each node can calculate the distance to all neighbors using time of arrival (TOA) principle. This algorithm estimates the location of neighbors in a local coordinate system referring to one of the node in the neighborhood as a reference. It is useable in places where GPS is not available.

III. ESTIMATION OF NODE'S LOCATION

Nodes equipped with a navigation system are capable of determining a path of their movement [2]. The path could be a route on a map, flight route, a pedestrian walking track, or any possible path. The knowledge of the path yields a set of valid points where the node is expected to pass along that path. Fig. 1 shows an illustration diagram for a mobile node carried by a vehicle travelling on a path; the path in this case is a map.

The mobile node that moves along a street in this case can navigate its way along the path and determines the coordinates of many points along this path using any digital mapping. The coordinates of these points are known, therefore, we predict that the mobile node will pass through these nodes but the moments of time when the mobile node is going to pass any of these points are unknown. Let point A with coordinate (x_0, y_0) be the starting point where the mobile node is lactated at time t_0. And let point B with coordinates (x_n, y_n) be the point where the mobile node is going to reach at time t_n.

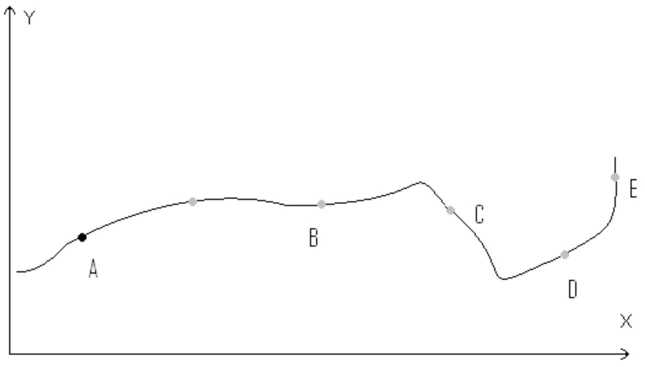

Fig. 1. A mobile node travelling along a path.

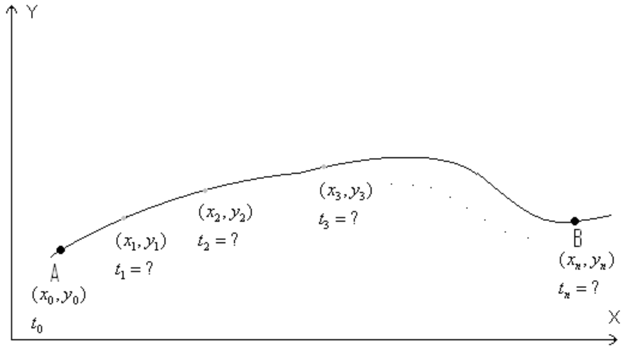

Fig. 2. The mobile node knows the coordinates of many discrete points between point A and point B.

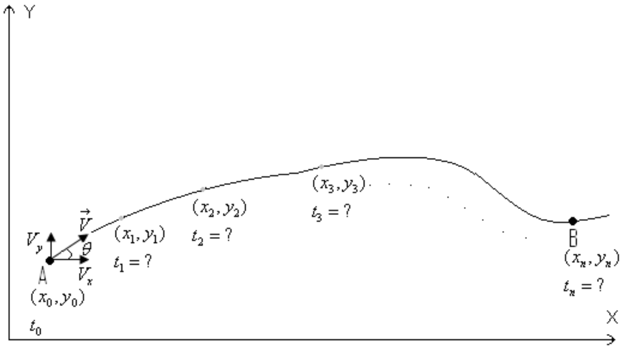

Fig. 3. A node moving at a constant speed $|\vec{V_0}|$, and initial direction θ.

We are interested in finding the equation for the locus formed by movement of nodes, with the following assumptions:

1) The relation between X and Y coordinates is a function; i.e. each X coordinate value has a unique Y value. $Y = F(X)$ or $X = F(Y)$.
2) Based on regular behaviour of the node, we assume the speed is constant during the estimation period (T), and the value of the speed is the instant value for the speed at point (x_0, y_0).

$$\vec{V_0} = \vec{V_x} + \vec{V_y} \tag{1}$$

$|\vec{V_0}|$ is constant on $[t_0, t_0 + T]$ while the X and Y components of the velocity are variables.

For simplicity, we assume a two-dimension movement $X - Y$, a three-dimension case is possible but with more calculations.

Let us look in more depth for the section of the path between point A and point B which represented by Fig. 2.

The X and Y coordinates for many points along the path between points A and B can be obtained and represented in n pairs of X and Y coordinates $(x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ Our aim is to use curve fitting technique, or polynomial regression as it mathematically known, to construct a function $y = f(x)$ of degree m that fits all points

$$(x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

. The polynomial is of the following form:

$$y_i = a_0 + a_1 x_i^1 + a_2 x_i^2 + \ldots + a_m x_i^m + \epsilon_i \tag{2}$$

Where $(i = 1, 2, \ldots, n)$. In matrix notation we have:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^m \\ 1 & x_2 & x_2^2 & \ldots & x_2^m \\ \vdots & & \ddots & & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^m \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \tag{3}$$

$Y = X\vec{a} + \epsilon$, where ϵ is a zero mean random error.

If we substitute the above points into (3) with $\epsilon = 0$, we will have a set of linear equations, these equations can be solved to find the values of the constants $a_0, a_1, a_2, \ldots, a_m$

$$\overset{\frown}{\vec{a}} = (X^T X)^{-1} X^T Y$$

If we substitute the value of the constants $a_0, a_1, a_2, \ldots, a_m$, into the general form of (2), we will get the polynomial that is the best curve to fit all points

$$(x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

$$y = a_0 + a_1 x^1 + a_2 x^2 + \ldots + a_m x^n \tag{4}$$

Equation (4) represents the change of node's location along the path between point A and point B. To find out how these location changes with time we consider the second assumption we have made, which states that the speed is constant during the period of estimation T.

Fig. 3 shows a mobile node moving at a constant speed along the path between point A and point B.

The node is moving along the path at constant speed $|\vec{V_0}|$ and variable direction θ, where θ is measured with respect to X-axis. The velocity of a mobile node describes the change in its speed or direction of movement or both. The best way to describe the velocity of an object is to find the rate of change of its position with time.

The relation between the X and Y coordinates is formed clearly in (4). The derivative of this equation can give us a view of how one variable is changing with respect to the other.

To find the relation between the X coordinate of node's location and time, let us start with the derivative of equation 4.

$$\tan\theta = \frac{dy}{dx} \tag{5}$$

This is also the slope of the tangent at each point

$$\Rightarrow \theta = \tan^{-1}\left(\frac{dy}{dx}\right) \tag{6}$$

Recall, $\vec{V_0} = \vec{V_x} + \vec{V_y}$.

Since $|\vec{V_0}|$ is constant on $[t_0, t_0 + T]$, therefore, the change is either by $\vec{V_x}$ or $\vec{V_y}$.

Since the speed at a certain direction is the change of position in that direction with respect to time; we can express

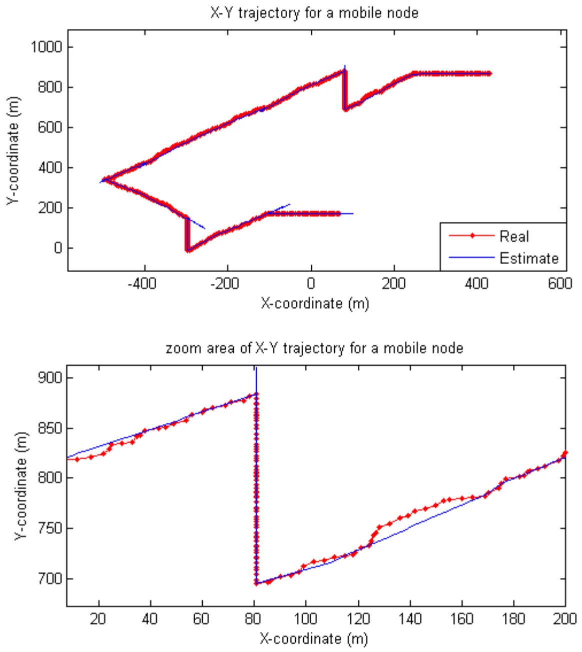

Fig. 4. A comparison between real and estimated trajectory with an estimation cycle $T = 30$ seconds, the lower part of the figure shows a magnified area for part of the plot for more details.

the X-component of the velocity by:

$$\vec{V_x} = \frac{dx}{dy}$$

We can also express the X coordinate of the location by:

$$\Rightarrow x = \int_{t_0}^{t} v_x dt \tag{7}$$

On $[t_0, t_0 + T]$

$$x = \int_{t_0}^{t} V_0 \cos\left(\tan^{-1}\frac{dy}{dx}\right) dt \tag{8}$$

On $[t_0, t_0 + T]$

This equation gives a relation between X and t during the period of estimation.

The values of Y can be obtained by substituting the values of X into the original coordination equation or by following the same steps for Y to get:

$$y = \int_{t_0}^{t} V_0 \sin\left(\tan^{-1}\frac{dy}{dx}\right) dt \tag{9}$$

within the duration $[t_0, t_0 + T]$.

The accuracy of the prediction model depends on the choice of the estimation time T. As the value of T decreases the model will converge to a line but the number of the estimated points will decrease. On the other hand, as T increases the number of future estimated points will increase, but with higher error. So the choice of T should be considered carefully.

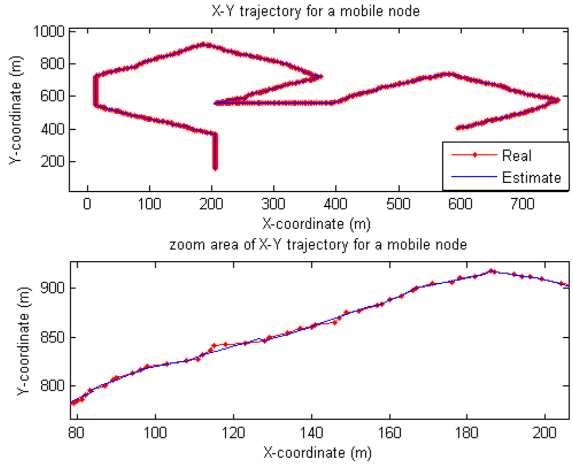

Fig. 5. A comparison between real and estimated trajectory with an estimation cycle T=5 seconds, the lower part of the figure shows a magnified area for part of the plot for more details.

IV. TOPOLOGY ESTIMATION

The model will estimate a set of future locations for the mobile node, and the expected time to traverse each estimated location. Each node will execute this estimation process every T seconds and then, disseminate the estimated values into the network; the dissemination could be addressed to all nodes in the networks or to the location servers. Table 1 shows the format of the update packet.

For topology based protocols, all nodes in the network will use the update packets to construct the global topology which will be an actual topology at the beginning of each estimation cycle. Then for each time step $t_1, t_2, \ldots, t_k$ nodes will use the estimated values of x and y coordinates to construct the estimated topology. During this period of time the nodes will not send updates for their locations and all nodes will use the estimated locations to construct the current topology.

For geographical protocols, nodes will use an estimated location for the destination stored in the location server.

V. EVALUATION OF SELF TRAJECTORY ESTIMATION PERFORMANCE

A computer simulation was carried out using MATLAB to evaluate the accuracy of the proposed estimation technique. We have set the area of the network as 1000m x 1000m. At the first phase of the simulation we generated a route from a random source to a random destination, and generated a movement for a node and record the coordinate of the node every second with its time stamp. The second phase of the algorithm was to estimate the trajectory of the node using the polynomial regression estimation technique, and

TABLE I
FORMAT OF UPDATE PACKET

Node ID	Control fields	x_0	x_1	…	x_k
		y_0	y_1	…	y_k
	…	t_0	t_1	…	$t_k = t_0 + T$

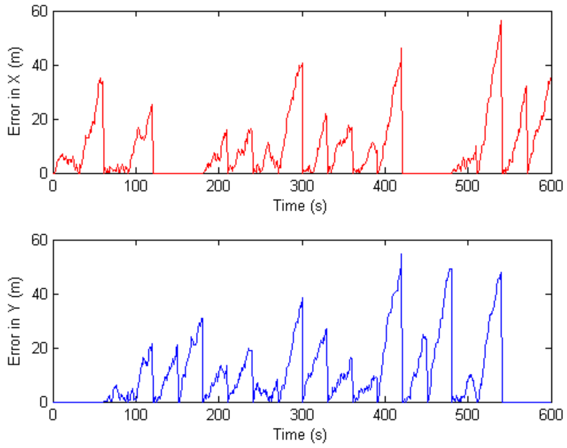

Fig. 6. The absolute error in both X and Y coordinates for $T = 30$ seconds.

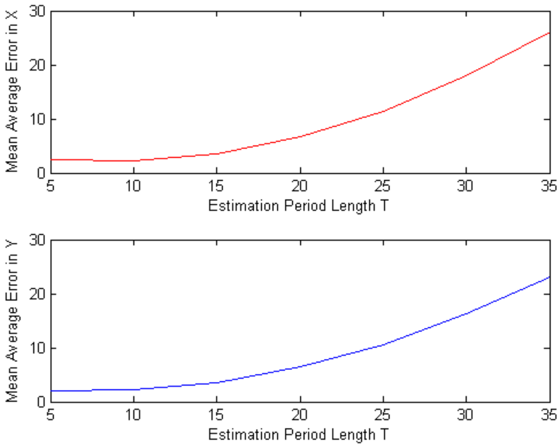

Fig. 8. Average absolute error as a Function of Estimation Cycle (T).

then a comparison between the estimated and the generated trajectories was carried out. This comparison is to test the accuracy of the algorithm and to measure the error during the estimation cycle. The effect of the estimation cycle (T) was studied as well by running the program for different values of T. The algorithm estimates a set of future locations along with the expected time for each estimated location. These points were used in estimating the topology of the network at different future times. The node travels on the path on an average speed of $60km/h$. Fig. 4 and 5 illustrate a comparison between real and estimated trajectory for a mobile node using two different estimation cycles ($T = 30$ seconds) and ($T = 5$ seconds).

We can noticed that the accuracy of the algorithm is higher when the estimation cycle is smaller. Fig. 6 and 7 illustrate the absolute error in both X and Y coordinates for two different values of T. From the above figures we can notice the following points:

1) The accuracy of the algorithm is inversely proportional to the estimation cycle T. The error on Fig. 6 where $T = 30$ seconds measures a maximum value of 50m for both X and Y, the value is reduced to less than $10m$ when the cycle $T = 5s$. Fig. 8 illustrates the relation between the average absolute error and the estimation cycle (T).
2) The value of the absolute error in X is zero when the path is parallel to the Y-axis because the estimation is carried only on Y coordinates. Also the error in Y coordinates is zero when the path is parallel to the X axis because the estimation is carried on X coordinates only.
3) The error varies from zero at the beginning of the estimation cycle because the node start with real measurements of location, speed and direction, then increases gradually to reach maximum value at the end of the estimation cycle (T).

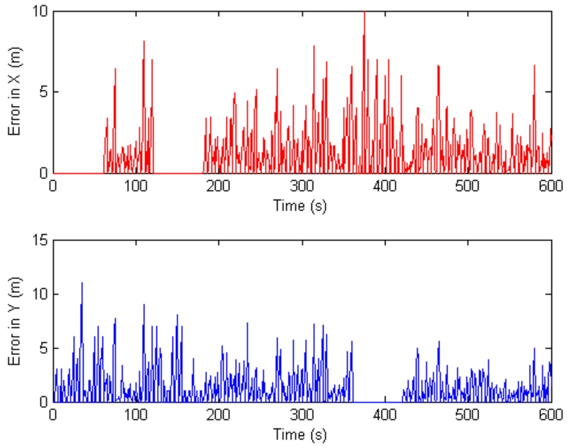

Fig. 7. The absolute error in both X and Y coordinates for $T = 5$ seconds.

VI. CONCLUSIONS AND FUTURE WORKS

We have demonstrated the polynomial regression trajectory estimation to use it in estimating the future topology of dynamic networks. We have developed a mathematical model based on polynomial regression to find the direct relation between X and Y coordinates for a mobile node. We have derived the relation between both X and Y coordinates and the time, for the period of estimation.

The algorithm gives good localization results, and it is capable of estimating the future location as a function of time which will be used in predicting the future topology of the network. The errors in the estimated values for X and Y coordinates are small compared to node's transmitting range in most type of mobile ad hoc networks nodes. Future work will include studying the factors that affect the choice of the optimal value for topology prediction cycle T, and to let nodes determine this value dynamically based on network and route conditions.

REFERENCES

[1] W. Su, S. J. Lee, and M. Gerla, "Mobility prediction in wireless networks," in *21st Century Military Communications Conference Proceedings*, vol. 1, 2000, pp. 491–495.

[2] E. D. Kaplan, *Understanding GPS: principles and applications*, 2nd ed. Boston: Artech House, 2006.

[3] M. Maroti, P. Volgyesi, S. Dora, B. Kusy, A. Nadas, A. Ledeczi, G. Balogh, and K. Molnar, "Radio interferometric geolocation," in *Proceedings of the 3rd international conference on Embedded networked sensor systems*, San Diego, California, USA, 2005.

[4] G. Dommety and R. Jain, "Potential networking applications of global positioning systems (GPS)," The Ohio State University TR-24, Tech. Rep., 1996.

[5] S. Basagni, I. Chlamtac, V. Syrotiuk, and B. Woodward, "A distance routing effect algorithm for mobility (DREAM)," *ACM/IEEE MOBICOM*, 1998.

[6] S. Woo and S. Singh, "Scalable routing protocol for ad hoc networks," *Wireless Networks*, vol. 7, pp. 513–529, 2001.

[7] Y. B. Ko and N. H. Vaidya, "Location-aided routing (LAR) in mobile ad hoc networks," *Wireless Networks*, vol. 6, pp. 307–321, 2000.

[8] B. Karp and H. T. Kung, "GPSR: greedy perimeter stateless routing for wireless networks," in *Proceedings of the 6th annual international conference on Mobile computing and networking*, Boston, Massachusetts, United States, 2000, pp. 243–254.

[9] H. Hurskainen, J. Raasakka, and J. Nurmi, "Specification of GNSS application for multiprocessor platform," in *International Symposium on System-on-Chip*, 2008, pp. 1–6.

[10] J. Fulton, "Global positioning system (GPS) technology an explanation of global navigation satellite system (GNSS)," Alabama Cooperative Extension System, Tech. Rep., 2009.

[11] S. Capkun, M. Hamdi, and J. P. Hubaux, "GPS-free positioning in mobile ad-hoc networks," in *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*, vol. 9, 2001.

[12] H. A. Seyin, K. Vassil, and D. Alex, "GPS-free node localization in mobile wireless sensor networks," in *Proceedings of the 5th ACM international workshop on Data engineering for wireless and mobile access*, Chicago, Illinois, USA, 2006.

[13] S. M. Das, H. Pucha, and Y. C. Hu, "Performance comparison of scalable location services for geographic ad hoc routing," in *24th Annual Joint Conference of the IEEE Computer and Communications Societies*, 2005, pp. 1228–1239.

[14] Z. J. Haas and B. Liang, "Ad hoc mobility management with uniform quorum systems," *IEEE/ACM Transactions on Networking*, vol. 7, pp. 228–240, 1999.

[15] M. Al-hattab and J. Agbinya, "Topology prediction and convergence for networks on mobile vehicles," in *Proceedings of the International Conference on Computer and Communication Engineering*, Kuala Lumpur, Malaysia, 2008.

[16] M. Al-Hattab and J. Agbinya, "Use of street maps to aid node localization in mobile wireless networks," in *International Symposium on Parallel and Distributed Processing with Applications*, 2008, pp. 471–476.

[17] A. Agarwal and S. R. Das, "Dead reckoning in mobile ad hoc networks," in *Wireless Communications and Networking*, vol. 3, 2003, pp. 1838–1843.

[18] V. Kumar and S. R. Das, "Performance of dead reckoning-based location service for mobile ad hoc networks: Research articles," *Wireless Communications & Mobile Computing*, vol. 4, pp. 189–202, 2004.

[19] N. Meghanathan, "Location prediction based routing protocol for mobile ad hoc networks," in *Global Telecommunications Conference*, 2008, pp. 1–5.

[20] T. Liu, P. Bahl, and I. Chlamtac, "Mobility modeling, location tracking, and trajectory prediction in wireless atm networks," *Journal on Selected Areas in Communications*, vol. 16, pp. 922–936, 1998.

[21] Z. R. Zaidi and B. L. Mark, "Mobility estimation for wireless networks based on an autoregressive model," in *Global Telecommunications Conference*, vol. 6, 2004, pp. 3405–3409.

[22] ——, "Real-time mobility tracking algorithms for cellular networks based on Kalman filtering," in *IEEE Transactions on Mobile Computing*, vol. 4, 2005, pp. 195–208.

[23] G. Eshel, "The yule walker equations for the AR coefficients," University of South Carolina.

[24] G. Welch and G. Bishop, "An introduction to the Kalman filter," University of North Carolina at Chapel Hill, Tech. Rep., 1995.

[25] K. Farkasa, T. Hossmann, F. Legendre, B. Plattner, and S. K. Das, "Link quality prediction in mesh networks," *Computer Communications*, vol. 31, pp. 1497–1512, 2008.