An Efficient Markov Decision Process Based Mobile Data Gathering Protocol for Wireless Sensor Networks†

Xin Fei
University of Ottawa
xfei@site.uottawa.ca

Azzedine Boukerche
University of Ottawa
boukerch@site.uottawa.ca

Richard Yu
Carleton University
richard_yu@carleton.ca

***Abstract*—Improving the performance of data gathering while saving the energy of network is one of the subjects of mobile data gathering protocols. In order to achieve the goal, the virtual force technique is introduced into the research where the data in the transmission queue and the distance between the sensor and the collector are simulated as forces. The collectors move on the direction of net force until the system reaches an equilibrium state. However, in the virtual force based data gathering system, the connectivity and communication quality have to be maintained for updating the changes of virtual forces. Otherwise, the loss of message will lead the system into a wrong equilibrium state and lower the performance of data gathering. In this paper, we formulate the moving process of data collectors as a Markov chain and determine the moving path using Markov decision process (MDP). The collectors move along the path defined by the optimized policy which is off-line computed and downloaded to collectors. Because the movement of collector no more relies on the online computed virtual force, the connectivity and communication quality of network have less effect on our MDP based protocol than the original virtual force based protocol. We compare the simulation results of our protocol, virtual force based protocol and the travel salesman problem (TSP) based protocol. The simulation results show that our protocol is resistant to connectivity error and works well when the communication error is applied to the system.**

***Index Terms*—data gathering, wireless sensor network, MDP, virtual force.**

I. Introduction

Three key performance measurements of data gathering system in the wireless sensor network are the data gathering rate, the report latency and the energy usage for delivering data. In order to save the energy of sensors while improving the performance of data gathering system, many mobile data gathering protocols have been proposed. In the protocols, researchers attach the sinks to mobile devices for collecting messages. Because the energy usage for communication propagates to the distance between collectors and sensors the moving path of collectors highly affected the performance of entire collection system. Virtual force based scheme is one of the widely used moving scheme for improving the efficiency of collector's movement. In the scheme, the distance between collector and sensor is simulated as a virtual force. Through computing the net force, the collector keeps moving to an optimized position where the energy usage for collecting data is minimized. The advantage of such an approach is that it does not require centralized control and localization. It makes virtual force approach distributive, robust and scale to very large networks. However, in order to measure the virtual forces, collectors have to periodically update the volume of data in the sensor's output queue and the distance between the sensor and the collector. The noise in the environment may introduce errors into the measurement of virtual force and further lead the system stopping at a wrong equilibrium state. Besides, the frequently updating process will cause the sensor consuming the energy quickly.

In order to avoid those problems, we propose our MDP based moving scheme where the movement of a mobile collector is modeled as a Markov chain and the selection of moving direction is decided by a Markov decision process (MDP). The paper is organized as follows: Section 2 introduces some related work and our motivations; Section 3 gives the virtual force model of energy-saving mobile data gathering; section 4 models the virtual force system by MDP; Section 5 proposes our MDP based data gathering protocol; Section 6 discusses the simulation environments and experimental results and is followed by the conclusion.

II. Related Work

Virtual force techniques are introduced into wireless sensor network as approximation algorithms for optimization in the path planing and auto deployment. The distance between mobile devices and obstacles is simulated as virtual repulsive and attractive forces which force the sensors to move away or towards each other. The system is considered to be optimized when the equilibrium state of system is reached.

Coverage-preserving deployment is the major area of virtual force model. In [1], the sensors and objects exert virtual repulsive force that pushes sensors away from the objects and also from each other so that their sensing areas do not overlap. The sensors will keep moving until equilibrium state is achieved; when repulsive and attractive forces are equal they end up canceling each other so that the full coverage is achieved. In changing environment, whether the change occurs

†This work is partially supported by Canada Research Chair Program, NSERC, Ontario Distinguished Researcher Award, Ontario EAR/PREA Award, and ORF/MRI funds.

periodically or sporadically, the network arrangement will be different with every change.

Poduri and Sukhatme[2] also apply the concept of virtual force to redeploy mobile sensors but with another constraint that each sensor needs to have at least K neighbors. Wang et al [3] propose three mobility algorithms to redistribute mobile sensors, namely VEC (VECtor-based), VOR (VORonoi-based) and Minimax. The VEC algorithm is a virtual force based algorithm and the VOR moves a sensor towards the farthest uncovered Voronoi vertex of this sensor. The Minimax algorithm is also based on Voronoi diagram and moves a sensor to the point where the distance to the farthest Voronoi vertex is minimized.

Because of the robust and distributed nature, virtual force model has been extended to many other applications such as mobile data gathering. In a large-scale wireless sensor network, a data gathering process is usually completed by the multi-hop data propagation. It consumes significant amount of energy, especially in the area near the sink. Because the energy consumption for data transmission propagates to the distance, obviously the most energy-efficient way for data collection is moving the collector as close as possible to sensors. Therefore, minimizing the energy consumption for collecting the data is equal to resolving a travel salesman problem (TSP), a problem of finding a shortest moving path to traverse the sensors. The problem of TSP based solution is that the sensors at the end of path suffer a long report delay. Many other algorithms have been proposed to resolve such problem such as SENMA (Sensor Networks with Mobile collectors) [4], AIMMS (Autonomous Intelligent Mobile Micro-server) system [5] and so on. In [6], we propose a virtual force based moving protocol for mobile data collection where data and distance are simulated as elastic forces. The net force points out the moving direction of mobile collectors. The problem of such algorithm is that it needs a stable communication environment. So, the control messages could arrive at the collector for computing the virtual force. The noise of environment may introduce errors into the measurement of virtual force and further lead the system stopping at a wrong equilibrium state. Besides, the frequently updating process will cause the sensors consuming the energy quickly.

III. ENERGY-AWARE VIRTUAL FORCE BASED DATA GATHERING

In this work, we consider an energy-saving mobile data gathering scenario presented in the [6] where sensors are randomly deployed into the field to surveillance the environment. In order to save the energy of sensors from relaying data to the sink, reduce the report latency and increase the data gathering rate, mobile collectors are deployed to gather data directly from sensors.

1) The Energy Model: We adapt the radio energy model presented in [7], [8], [9] where both Friss free space model and two-ray ground model are used. A crossover distance $d_{crossover}$ is defined as follows:

$$d_{crossover} = \frac{4\pi\sqrt{L}h_r h_t}{\lambda} \quad (1)$$

where L is the system loss factor; h_t and h_r are the height of the transmitting and receiving antennas above ground; λ is the wavelength of the carrier signal (for consistency we adopt the notations of NS2). When the distance between the sender and receiver is lower than $d_{crossover}$, the Friss free space model is used; otherwise, two-ray ground propagation model will be adapted.

If the propagation loss of short and long distance is modeled as variables that are inversely proportional to d^2 and d^4 respectively, the energy usage for transmitting m bits through distance d can be formulated as:

$$\begin{aligned} E_T &= mE_{elec} + mE_{ampl}d^2 \quad when \quad d \le d_{cross-over} \\ E_T &= mE_{elec} + mE_{ampl}d^4 \quad when \quad d \ge d_{cross-over} \end{aligned} \quad (2)$$

where E_{elec} is the energy cost of the modules such as coding and modulation module that process data before transmission. Assuming that E_{elec} is a constant small value for every sensor, the energy usage of a transmission E_T is only related to the distance d and the energy consumption of the amplifier E_{ampl}. Based on the equation 2, the total energy consumption E_{total} for transmitting data to sink through multi-hop propagation can be formalized as:

$$E_{total} = \sum_{i=1}^{n} \sum_{j=1}^{hops} m_i E_{ampl} d_j^2 \quad (3)$$

2) Virtual Force Model for Mobile Data Gathering: In order to maximize the energy efficiency, mobile data gathering protocol has to minimize E_{total} in the formulation 3. In other words, the moving protocol for mobile collectors has to reduce the hops and distance of transmission for the entire network by moving collectors toward those sensors with more data than the others. Comparing to the energy usage of amplifier

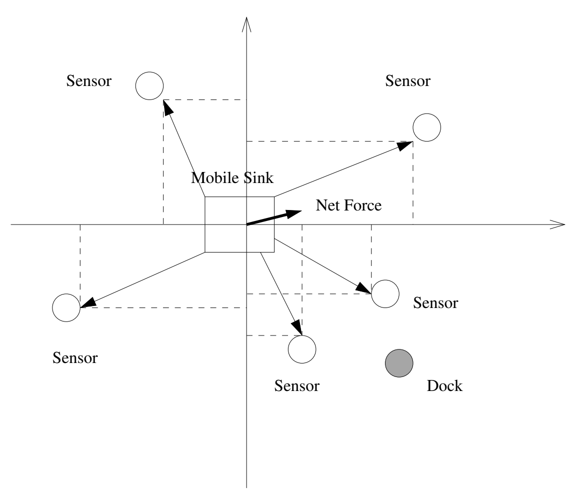

Fig. 1. Net Force

E_{ampl}, E_{elec} is very small. If we ignore the energy usage of other modules (E_{elec}) and only consider the transmission under free space model, the energy usage for transmitting m bits is $mE_{ampl}d^2$. Therefore, based on hooke's law, the energy used for transmitting m bits over distance d is very similar to the work done by an elastic force where the spring constant K is determined by mE_{ampl} as following.

$$\begin{aligned} &F = Kd \\ &W = K\int_0^d d^2 = P_t = mE_{ampl}d^2 \\ &K \approx 2mE_{ampl} \end{aligned} \tag{4}$$

As shown in Figure 1, after simulating the energy usage for transmitting data as an elastic force, the problem of minimizing $\sum mE_{ampl}d^2$ is equal to find an equilibration point where the net force is zero. The net force determines the direction of movement.

For a static optimization problem, such as coverage-awareness deployment, the equilibrium state does not change after it is discovered. However, for the mobile data gathering problem, because the events in the network are randomly distributed the equilibration points will change frequently. Since the speed of collector is low, when the working area of collector is large the equilibration point is always expired before collectors arrive at it.

3) The Number of Mobile Collectors: The virtual force model shows how the collectors move in the environment. However, it does not mention how to determine the number of mobile collectors. For a given area, what are the relationships among moving speed, waiting time and the number of collectors? Is there any upper bound for the number of collectors?

We first consider a scenario that when a sensor wants to send data there is at least one collector which could arrive at the sensor in time t. If the speed of collector is V and the acceptable waiting time for data delivery is t then we have an extend coverage range $r' = V*t+r$ where r is the transmission range of collectors. Assuming that distribution of sample X follows normal distribution, the probability that a sample x falls in a collector's cover range is proportional to the area of sensing disk. The probability function of $dist(x,c) < r'$ is:

$$F(r') = \begin{cases} 0, & r < 0; \\ \frac{r'^2}{R^2} & 0 \le r < R; \\ 1 & r' \ge R \end{cases}$$

where R is the radius of the area.

When m collectors are applied to the environment the probability that the data of sensor could be gathered in time t is:

$$p = 1 - (1 - f(r'))^m$$

The Figure 2 shows the function p for an area where $R = 100$. When the cover range is equal to the radius of area only 1 mobile could cover entire area. When the cover range is 30 we need at least 20 collectors to keep a high probability for collecting data efficiently.

The result shown in Figure 2 is assuming that the collector is uniformly distributed in the area. However, the question is that do we need so many collectors. A simple upper bound of m is that m collectors could cover the entire area. Kershner[10]

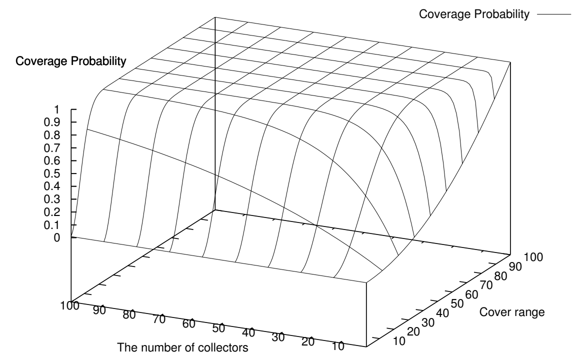

Fig. 2. Coverage Probability

has pointed out that the lowest redundancy rate for covering a rectangle by circles is $2\pi/\sqrt{27}$. Therefore, for a 1-cover node distribution the number of deployed collectors can be found as following.

$$Limit\left(n\pi r^2\right) \to \frac{2\pi}{\sqrt{27}}\bar{M}(1 - cover)$$

Based on above function, for a sparse distribution of sensors, a upper bound of m could be derived from looking for the polygons enclosing all the sensors.

IV. MDP Modeling of Virtual Force Model

In the virtual force based moving scheme, mobile collectors move to the optimal position determined in each round. If we consider the virtual force based moving plan as a periodical decision system, then the mobile collector makes a decision on the moving direction and speed at each stage. However, the virtual force model gives a locally optimized result because the collector can only collect the information of forces in its coverage area. Besides, the optimal positions are hard to be reached for the limitation of mobility of mobile device. In order to find an optimal moving plan for a long run, in this section, we will resolve the path plan problem of mobile data gathering using MDP.

Before we start modeling the data gathering problem, we simplify the data gathering problem presented in the last section. We assume that every sensor s_i has a probability P_i to generate a $xbit$ packet at each time. The generation probabilities of sensor are independent with each other.

In the rest of this section, we will first introduce the Markov decision process (MDP) and then model the moving prediction problem as a Markov decision process. The optimization goal of MDP is to find an optimized policy composed of the optimal path for traveling. In order to concentrate on the path planing problem, in our MDP model we will ignore the effect of channel constraint and time constraint by assuming the existence of the TDMA and large queues.

A. MDP

An MDP (Markov Decision Process) is a 4-tuple(S, A, T, R). S is a finite set of states, in one of which the world exists. A is a set of actions that may be executed at any state. T is the transition probability that defines how the state changes when an action is executed. It is denoted as $T(s, a, s')$. The probability of moving a state to another state is dependent only on the current state (the Markov property). $R(s, a)$ is the reward function for performing action a when at state s. The optimization purpose is finding a policy which is defined as a function that determines an action for every state $s \in S$. The quality of a policy is the expected sum of future rewards. Future rewards are discounted by a discount factor β to ensure that the expected sum of rewards converges to a finite value, i.e., a reward obtained at t steps in the future is β^t, $0 < \beta < 1$. The value of a state s under policy π, is the expected sum of rewards obtained by following the policy π from s.

B. States and Transitions of System

The mobile collector is modeled by a variable c whose states is the location of collector (x, y). At each stage, the mobile collector will determine the moving direction θ and the speed V . Therefore, the action is a tuple (θ, V). In order to simplify the problem, as shown in Figure 3 we convert the continuous state Markov chain into a discrete state Markov chain by dividing the space into grids. The diameter of grids is V_{max}, for example, it is $6m/s$ in our simulation. Because the speed of collector is determined by its weight and the friction of ground surface, we can assume that for a collector its speed V will maintain same in a small area, such as the area from one grid to its neighbor grids. In order to further simplify the problem, we remove the V from action space and reflect it by the one-step transition probability from one grid to another when the θ is selected. Therefore, the moving function is replaced by nine possible grid assignments where one of the nine grids is selected as the destination of the collector at each stage. The one step transition probability from a cell to another cell is zero except the eight neighboring cells and itself.

C. The Reward Model

The goal of mobile data gathering protocol is to minimize the energy usage for collection, $\sum_{i=1}^{n} m_i E_{ampl} d(s,i)^2$, where $d(s, i)$ is the distance between collector and sensor i, and maximize the volume of collected data, $\sum_{i=1}^{n} m_i$. Therefore, we define the reward function for MDP as following:

$$R(s, a, s') = \left(\sum_{i=1}^{n} m_i \right) - \varepsilon \left(\sum_{i=1}^{n} m_i E_{ampl} d(s,i)^2 \right) \quad (5)$$

$$R(s, a, s') = \sum_{i=1}^{n} m_i \left(1 - \varepsilon \left(E_{ampl} d(s,i)^2 \right) \right)$$

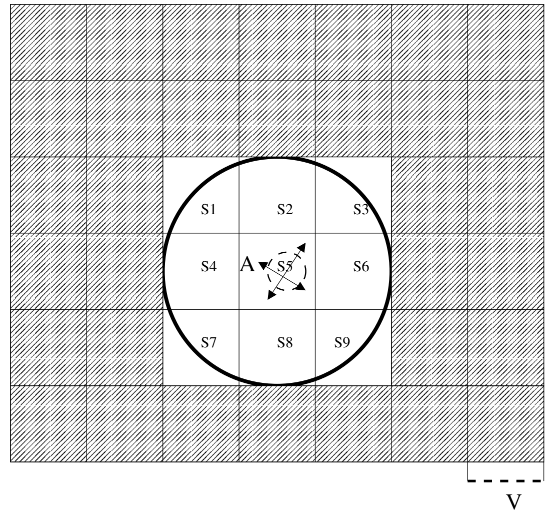

Fig. 3. Makrov Chain for Moving

where ε is a trade-off factor. The object of optimization is to maximize the $R(s, a, s')$.

In order to avoid the effect of dimension, we normalize the data to $[0, 1]$ and revise the reward function as following

$$R'(s, a, s') = \left(\sum_{i=1}^{n} f(m_i) \right) - \varepsilon \left(\sum_{i=1}^{n} f(d(s,i)) \right) \quad (6)$$

$$d_i = d(s, i)$$

$$f(m_i) = \begin{cases} \frac{m_i - m_{\min}}{m_{\max} - m_{\min}} = \frac{m_i}{Size_q} & if \;\; d_i \le r \\ 0 & if \;\; d_i > r \end{cases}$$

$$f(d_i) = \begin{cases} \frac{m_i d_i - m_{\min} d_{\min}}{m_{\max} d_{\max} - m_{\min} d_{\min}} = \frac{m_i d_i}{Size_q r} & if \quad d_i \le r \\ 0 & if \;\; d_i > r \end{cases}$$

where d_{max} is equal to the communication range r, $d_{min} = 0, m_{min} = 0$ and m_{max} is the size of queue $Size_q$. Since only the sensors in the coverage will exert the force to the collector, $f(m)$ and $f(d(s, i))$ are equal to zero when $d(s, i) > r$.

V. The Mobile Data Gathering Algorithm Using MDP Based Path Planning

A mobile data gathering algorithm using MDP based path plan includes two phases. In the off-line phase, based on the history data the sink computes the optimal policy and sends the policy to the collector to start the online phase. In the online phase, the collector periodically detects its current location by GPS and selects an action from policy to move. After t_m seconds, which is $1s$ in our simulation, the collector stops and broadcasts the REQ message to request sensors in its range uploading data for t_c seconds. If there is no data arriving in t_{idle} seconds, the collector detects its location and moves to next position.

In the virtual force algorithm, in order to determine the moving speed and direction, collector has to broadcast a query message to every sensor in its range. Sensors reply such query messages to announce their state. Therefore, the sensor network costs $\sum_{i=1}^{k} (ES_i + ER_i)$ energy at each round for the control messages, where ES and ER are the energy used for sending and receiving a control message respectively. Correspondingly, in the MDP based algorithm, collector could determine its state by GPS and select the moving direction from the optimal policy. In each round, the only control messages are the messages used to ask sensors to upload data. The total energy cost of such messages is $\sum_{i=1}^{k}(ER_i)$. Comparing to the virtual force based path planning, the MDP based algorithm doesn't have to collect the information of virtual force to determine the next move of collector. The collector could move to the next position and start collecting data immediately. Therefore, the MDP based solution has shorter report latency than virtual force based solution. Considering the limited energy of sensors, using less control messages could also benefit the lifetime of sensors.

The pseudo code of mobile data gathering algorithm using our MDP based path plan is shown as Algorithm 1.

Algorithm 1 The Mobile Data Gathering Algorithm

```
procedure OFF-LINE PHASE(DataLog)                ▷
    Generate the transition matrix;
    Compute the policy;
    Download the policy to collector;
end procedure
procedure ONLINE-PHASE(π)              ▷ π is the policy
    Broadcast the REQ to neighboring sensor;
    Determine current location;
    Move to the next location;
    Start Data Gathering;
    Repeat the Online-Phase;
end procedure
```

VI. SIMULATION EXPERIMENTS

In this section, we illustrate the performance of the proposed scheme by simulation experiments. The simulation results of our protocol are compared to the result of virtual-force algorithm and TSP based data gathering algorithm. The performance of schemes is evaluated in terms of data gathering rate, report latency and message complexity. The simulation environment is defined at following section.

A. The Simulation Environment

In order to simplify the simulation experiment scenarios, we consider a sensor network with $10 - 80$ randomly deployed sensors in a $100m * 100m$ area. The radio range of the sensor is $20m$. The moving speed V_{max} of the collector is $6m/s$. The data generation rate of each sensor is independent and follows a poisson distribution. The collector changes its moving destination every second. The rest of simulation parameters are shown as Table I.

B. Performance Results

In this section we compare the performance of algorithms under different scenarios. The result shown in this section is based on 40 runs for each scenario.

TABLE I
SIMULATION PARAMETERS

The size of space	100meter*100meter
The number of sensors	10-80
Radio range of sensor	20 meter
Moving speed of collector	6 meter/s
Size of packet	1k bytes
Packet Generation Rate	1/sec
Bandwidth	2M bps
T_m	1 sec
T_c	1 sec

1) Average Report Latency: The average report latency is used to measure the time duration from data generation to data collection. According to the transmission time model of NS2, the transmission time for delivering $1k$ bytes data over a $2M$ bps channel is around $32ms$. Therefore, the report latency shown in the Figure 4 is mainly contributed by the waiting time for the arrival of collectors. In the figure, the MDP based algorithm has a better performance than both virtual force based algorithm and TSP based data gathering algorithm. The TSP based algorithm has the worst performance because it only considers the effect of distance and ignores the data generation rate of each sensor. The reason for the virtual force based algorithm losing in the competition is that it only concernes the sensors in the range of the collector while the policy generated by MDP maximizes the total reward of entire running time.

When the density of deployment increases, the performance of three algorithms gets close especially for the TSP based algorithm. The reason of such phenomenon is that the average data generation rate of each grid becomes closer when the density of deployment increases. At the same collection period, in the three algorithms, the volumes of data collected by the collector are almost the same. Therefore, the travel of the collector does not benefit to the average report latency.

In order to show the performance under environment with noise, we add an receiving probability to the receiving model of ns2. Therefore, even the energy level of received message is higher than the receiving threshold, the message still could be dropped. The Figure 5 shows the performance changes when the receiving probability is applied to a sparse deployment (30 sensors are deployed). The report latency becomes longer because the transmission failure will cause data retransmission. The performance of MDP based algorithm is better than the virtual force based algorithm since the former algorithm does not rely on the online computing to get the net force. The loss of control message has less affection on its performance.

2) Data Gathering Rate: The data gathering rate is the average number of data collected by the algorithm during a running time. The Figure 6 shows the data gathering rate in an ideal environment where communication errors are not applied. In the same collection time, the MDP based algorithm gives the best result while the virtual force based algorithm ranks the second. Because TSP based algorithm only considers the travel distance, it shows the worst average data gathering

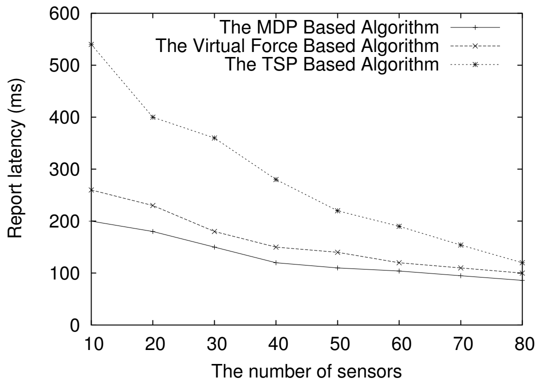

Fig. 4. The Report Latency Without Error

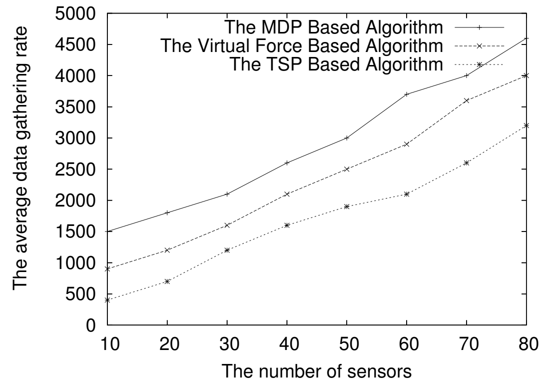

Fig. 6. The Data Gathering Rate Without Error

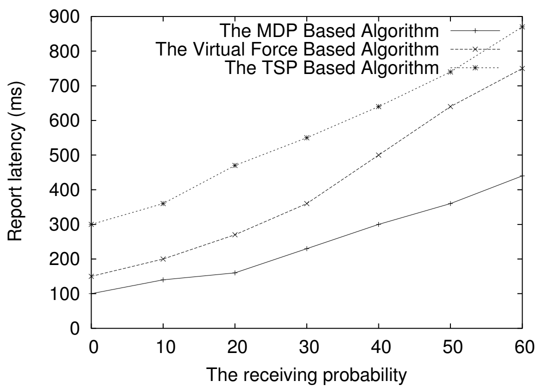

Fig. 5. The Report Latency Rate With Error

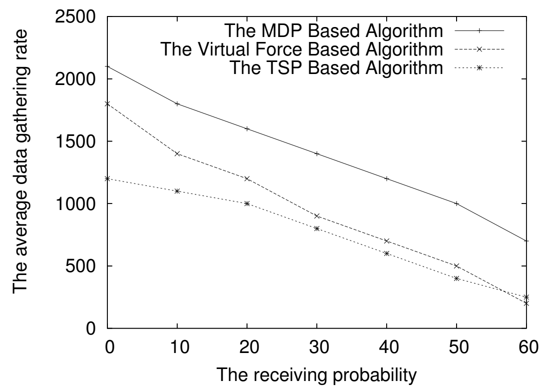

Fig. 7. The Data Gathering Rate With Error

rate in all scenarios. When the receiving probability is applied, the performance of virtual force based algorithm drops faster than MDP based algorithm as shown in Figure 7. When the receiving probability is over 54 percent the performance of virtual force based algorithm is even worse than the TSP algorithm. The reason is that the loss of control messages leads to the incorrect net force.

VII. Conclusion

In this paper, we have presented a MDP based virtual force algorithm for mobile data gathering. The movement of the collector is modeled as a Markov chain. Instead of computing net force on-line, the moving path of the collector is determined by an optimized policy which is generated from a Markov decision process. The collector can move in the area without control messages. We compared the performance of proposed scheme with general virtual force based data gathering algorithm and TSP based data gathering algorithm. The simulation results show that the proposed scheme is suitable for working in a data gathering scenario where the connectivity and communication quality can not be guaranteed. Currently the presented algorithm only support a single collector. However, multiple collectors could be managed by modeling the multi-collector data gathering problem as a multi-armed bandit problem.

References

[1] A. Howard, M. J. Mataric, and G. S. Sukhatme, "Mobile sensor network deployment using potential fields: A distributed, scalable solution to the area coverage problem," in *Distrib Auton Robot Syst 5*, 2002, pp. 299–308.

[2] S. Poduri and G. Sukhatme, "Constrained coverage for mobile sensor networks," in *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, vol. 1, apr. 2004, pp. 165 – 171.

[3] G. Wang, G. Cao, and T. La Porta, "Movement-assisted sensor deployment," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 4, mar. 2004, pp. 2469 – 2479 vol.4.

[4] L. Tong, Q. Zhao, and S. Adireddy, "Sensor networks with mobile agents," in *Military Communications Conference, 2003. MILCOM 2003. IEEE*, vol. 1, oct. 2003, pp. 688 – 693 Vol.1.

[5] W. J. K. M. B. S. G. J. P. A. Kansal, M. Rahimi and D. Estrin, "Controlled mobility for sustainable wireless networks," in *SECON 2004*. ACM, 2004.

[6] A. Boukerche and X. Fei, "Adaptive data-gathering protocols with mobile collectors for vehicular ad hoc and sensor networks," in *WiMob*, 2008, pp. 7–12.

[7] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols," in *MobiCom '98: Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking*. New York, NY, USA: ACM, 1998, pp. 85–97.

[8] T. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.

[9] W. B. Heinzelman, "Application-specific protocol architectures for wireless networks," Ph.D. dissertation, 2000, supervisor-Chandrakasan, Anantha P. and Supervisor-Balakrishnan, Hari.

[10] R. Kershner, "The number of circles covering a set," in *Amer. J. Math. 61*, 1939, pp. 665–671. 9.