Analytical Models of Short-Message Reliability in Mobile Wireless Networks

Debessay Fesehaye Kassa†, Klara Nahrstedt† and Guijun Wang‡
† Dept. of Computer Science, University of Illinois at Urbana-Champaign, USA. {dkassa2,klara}@illinois.edu
‡ Boeing Research & Technology, WA, USA. guijun.wang@boeing.com

***Abstract*—Applications like Twitter which use chat-like short messaging systems (SMS) have been widely used in public, political, military, emergency, humanitarian and other fields. Such applications usually involve servers (controllers) which control and forward messages from a sending client to a receiving client. The use of mobile wireless networks for such messaging systems has been increasing at a fast pace. To cope with this increase, there need to be efficient communication protocols and algorithms. To design such protocols and algorithms requires extensive analysis and understanding of the behavior of the communicating nodes under a given mobility scenario. One of the good metrics to understand the performance of such protocols is the reliability of message delivery.**

In this paper we present analytical models of the average reliability of short (chat-like) message delivery in mobile wireless networks as a multivariate function of the transmission range, movement area dimensions, number of servers (base stations) and message deadline (lifetime) under moderate realistic assumptions which can be easily relaxed and extended.

Simulation results show that our analytical models give very good estimation of the average reliability of message delivery.

I. Introduction

With the growth of wireless and mobile network technologies many interesting applications have emerged. Some of the most common and popular applications such as Twitter [4] and SMS [12] are used to transfer short chat-like messages. Such applications have proven to be very useful in various public, political, military and emergency fields where transfer of big files is either not possible, costly, or unnecessary.

Some of the nodes in such chat-like short messaging systems are servers (controllers) which are responsible for the control and transfer of messages from the sender clients to the receiver clients. Even though there are many messaging and presence protocols for short messaging systems, in this paper we focus on protocols like the Extensible Messaging and Presence Protocol (XMPP) [7]. In the XMPP protocol, all senders and receivers register at their respective controller (server) with their names or IDs. A server broadcasts presence information of the subscribed clients to all clients and other servers. A client creates its roster by getting subscription approval from other clients. The client can then send messages to its respective peer. A moving client can subscribe to the nearest server (smart selection) which can route its messages to the intended receiver or from which it receives its messages.

With the growth of such short-messaging applications [8], [12], understanding their performance in a given network scenario becomes crucial to design efficient protocols and to efficiently dimension the network under which they operate. An important performance metric for such applications is the average reliability of message delivery from one client node to another client node via some controller nodes (servers) in given network.

In mobile wireless networks, the value of the average message reliability becomes low either because the nodes are moving outside the coverage area or due to congestion and interference which results in increased delay and packet losses. In chat-like short messaging systems the decrease in average message reliability is mainly due to nodes moving outside the coverage area or due to clients moving far away from their servers. This is because the messages generated by such systems are too short to congest the underlying dedicated physical or overlay network. Besides, the message processing and forwarding delays at the servers are negligible in such short messaging systems. So in this paper we focus on the reliability of short messaging systems where the reliability loss is mainly due to nodes moving outside the coverage area.

Performance metrics such as reliability can be studied using simulation or analytical models. Analytical models have the advantage that they are faster than simulation. Analytical models can also give detailed insight to the applications and protocols studied via the closed-form expressions which show how reliability evolves as a function of each parameter.

In this paper we present analytical models for the reliability of short message exchange between nodes via servers in mobile wireless networks as a multivariate function of the transmission range, movement area dimensions, number of servers (base stations or access points) and message deadline under moderate assumptions which can easily be relaxed and extended. In our analysis, messages are buffered at all servers until their deadline expires or until the intended mobile receiver(s) gets them from any of the servers before they expire. Any other message replication scheme can be used with our modeling approach. The fact that some (multimedia) applications tolerate some packet losses (reliability loss) makes our reliability models specially useful. This is because our analytical models give closed form expressions of what parameters give the average desired reliability values.

The rest of the paper is organized in such a way that we first present some related works, how they differ from our scheme and why they are not sufficient in Section II. Following this we describe our analytical models of reliability in Section III. We then present some numerical results and a summary in Sections IV and V.

II. Related Work

There have been some works in the literature on the reliability of message delivery. Models of reliability in distributed publish-subscribe systems have been presented in [5], [6]. In this study the authors assumed that publishers (senders) are not mobile. They also assumed a specific network topology with fixed number of servers, transmission range and deadline parameters. Hence they didn't give closed form expressions of reliability as a function of the number of servers, transmission range, area dimensions and deadline parameters which are the main factors which affect the reliability of message delivery. This is specially useful to plan on how many servers with a specific transmission range are needed to obtain a specific reliability value in a given area and mobility scenario. A design and performance analysis of soft hand-off scheme for CDMA cellular systems was presented in [11]. The authors aimed to decrease both the number of dropped hand-off calls and the number of blocked calls without degrading the quality of communication service and the soft hand-off process. The authors also have given continuous-time Markov chain models to analyze their design. Their paper focused more on hand-off and a channel resource shortage condition.

III. Our Analytical Models of Reliability

In this section we present simple and advanced analytical models of reliability. We first discuss the notations and assumptions we use in our modeling. We next present some closed form expressions for the number m of servers needed to give full wireless coverage in the movement area and for the expected maximum tolerable delay (message life time) τ. The value of τ is the expected maximum length of time before a moving client gets its message. We finally present the simple and advanced analytical models for the average reliability of short message delivery.

A. Notations

We use the following notations to derive the closed form expression for the average reliability. We use the terms server, base station, and access point interchangeably. This is because the servers (message) controllers are more like access points or base stations in that the mobile senders cannot send and the mobile receivers cannot receive messages unless they are in range of the servers which are connected either using reliable wireless, wired or overlay network.

TABLE I
Simulation Parameters

Parameters	Description
A	area within which nodes move
n	number of clients which move in the area
m	number of servers needed to get full coverage
m_a	number of available servers
t	transmission range of a server
R	desired reliability
d	deadline of message delivery
c	speed of light
n_c	number of clients connected
l	length in meters of the rectangular area
w	width in meters of the rectangular area
v	speed of a node (client)

B. Assumptions

Unless and otherwise specified, in this paper we make the following assumptions which can be relaxed.

- The n clients are uniformly distributed (located) in the area A.
- The m servers (BS) should be uniformly placed in the area A to get full coverage.
- Each server is placed at the center of each square cell as shown in Figure 1.
- The servers are placed in such a way that the overall coverage of the movement area is maximized.
- The movement of the nodes is uniformly distributed in the area.
- The area within which the nodes move and the servers are placed is considered to be rectangular. For other shapes of the area, the rectangle circumscribing the shape is considered to obtain the worst case reliability value.
- All messaged from the mobile senders are buffered at all static servers until they are delivered to their respective mobile receivers or until they expire.

We next present the algorithms and derivations for the reliability of message delivery for short messaging systems in mobile wireless networks.

C. Number of Servers for Full Area Coverage

In this section we present some expressions for the total number of servers needed to give full wireless communication coverage for all mobile nodes in the movement area. Let us consider the mobility area as given in Figure 1 with n clients which want to communicate with each other. We then use the following simple scheme to derive the relationships between the short message transfer reliability R and the variables n_s, t, l, w and d shown in Table III-D.

- We first surround the general movement area for the mobile network with a rectangle as shown in Figure 1.

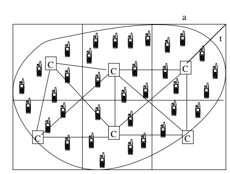


Fig. 1. The mobility area with n mobile nodes and m servers (controllers C)

- Placing all m servers in the area we partition the movement area into smaller (equal) square areas where each square is covered by one server. For cases where the rectangular area cannot be divided into an integral number of square cells, a fraction of some of the squares in the last column and/or row of the rectangular area can be considered.
- Taking one square, as shown in Table II with the dimension a, we derive a as a function of the transmission range t as shown in Equation 1.

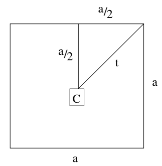


TABLE II
SQUARE AREA COVERED BY ONE SERVER

$$\begin{aligned}(\frac{a}{2})^2 + (\frac{a}{2})^2 &= t^2 \\ 2(\frac{a}{2})^2 &= t^2 \\ \Rightarrow a &= \hat{a}(t) \\ &= \sqrt{2}t.\end{aligned} \quad (1)$$

There are m such squares. Hence we have

$$A = ma^2 = 2mt^2. \quad (2)$$

Which implies that the total number of servers m with transmission range t needed to get full coverage assuming that is

$$m = \hat{m}(A, t) = \frac{A}{2t^2}. \quad (3)$$

- The nodes in the given area with the furthest distance are the once at two opposite corners of the rectangle. If the rectangular area has a length of l and a width of w, the furthest distance between two nodes is $\sqrt{w^2 + l^2}$. Denoting the speed of light with c, the maximum total queuing, transmission and processing delay with D, we get a reliability R of 1.0 (full reliability) of message delivery with m servers if the deadline

$$d > \frac{\sqrt{w^2 + l^2}}{c} + D. \quad (4)$$

We call this a deadline constraint. For short chat-like message delivery which is the main focus of this paper the value of D can be negligible.

- With n mobile clients sending and receiving messages and with m servers having a range of t, a full reliability of short message delivery can be achieved. To have a full coverage, the number of squares (servers) l_m in the horizontal (length) dimension of the rectangular area where nodes can move is given by

$$l_m = \hat{l}_m(l, t) = \lceil \frac{l}{a} \rceil = \lceil \frac{l}{\sqrt{2}t} \rceil = \lceil \frac{\sqrt{2}l}{2t} \rceil \quad (5)$$

and the number w_m of squares along the vertical (width) dimension is given by

$$w_m = \hat{w}_m(w, t) = \lceil \frac{w}{a} \rceil = \lceil \frac{w}{\sqrt{2}t} \rceil = \lceil \frac{\sqrt{2}w}{2t} \rceil. \quad (6)$$

Now the total number of servers in this square layout needed for the full coverage is

$$m = \hat{m}(l, w, t) = l_m \times w_m. \quad (7)$$

It should be noted that Equation 3 assumes that $\frac{l}{a}$ and $\frac{w}{a}$ are integers.

If we have k less servers than m with a range of t, then the messages sent by the clients at the k cells at a given time t will not be transmitted, as the clients are out-of range of any server. However the messages sent by the other clients which are in range of a server are received before they expire (before the deadline d) if there are sufficient number of servers and if the destination clients which are out-of range of a server move sufficiently fast. This is because all messages transmitted by the nodes are buffered at all servers before they expire or before they are consumed by their respective receiver.

D. Expected Maximum Message Delivery Delay

Under this buffering assumption discussed above, the message life time or deadline τ depends on the number n_s of servers, transmission range t, node speed v and coverage area dimensions l and w and can be expressed by a function $\hat{\tau}$ as follows. Let's assume that the decision of how many servers to place in a given region of a coverage area is made based solely on how much of the area can be best covered without giving special preference to some regions of the coverage area. If some preferences need to be given to some regions, then the whole coverage area can be partitioned into sub areas each of which has no preference over its regions. To achieve this a simple relationship can be drawn on how many servers to initially place in each sub area based on some weights or user specified policies. Our scheme can then be applied to each sub area and extended to the coverage area as a whole.

To derive the deadline or message lifetime τ, we will first find the maximum possible distance between a client and its nearest server and then use some work in the literature [1] to get the expected transmission time (deadline τ) from the uncovered cell to the covered cell. We call this variable a radius r of circle centered at the client and passing through one of the nearest servers as also used in [1].

As can be seen from Figure 2(a), with the assumption that there is no preference in any specific region of the coverage area, if the number of missing servers k is less than or equal to half of the total number m of the servers required to give a full coverage of the area, then the maximum distance or radius r between a client and the nearest server (BS) is $a = \sqrt{2}t$. If $k > m/2$, the value of r depends on whether or not l_m and w_m are both even, one of them even or both odd numbers. So we have two cases.

Case A: l_m and w_m are odd or one of them is even
This case is shown in Figures 2(b) and 2(c). In this case if only the outer most round of the cells is uncovered, $r = 2t$, which is the distance from one corner of the uncovered cell to the nearest corner of the nearest covered cell in the next inner round of the coverage area. So as long as only

$$k \leq \lfloor \frac{m}{2} \rfloor + \frac{2(l_m + w_m) - 2(2)}{2}$$

servers are missing $r = 2t$. Here $2(l_m + w_m) - 2(2)$ is the total number of cells in the outer most round of the rectangular coverage area. In general, if the i outer most rounds of cells are uncovered, $r = i \times 2t = 2it$. For ρ such rounds, $1 \le i \le \rho$ where

$$\rho = \lfloor min\left(\frac{l_m}{2}, \frac{w_m}{2}\right)\rfloor.$$

Hence for this case denoting

$$\begin{aligned} S_{i-1} &= \sum_{j}^{i-1}\left(\frac{2(l_m+w_m)-2(2)(2j-1)}{2}\right) \\ S_i &= \sum_{j}^{i}\left(\frac{2(l_m+w_m)-2(2)(2j-1)}{2}\right), \end{aligned}$$

for $1 \le i \le \rho$ we have

$$r = \begin{cases} \sqrt{2}t & \text{if } 0 < k \le \lfloor\frac{m}{2}\rfloor \\ 2t & \text{if } \lfloor\frac{m}{2}\rfloor < k \le \lfloor\frac{m}{2}\rfloor + S_1 \\ 2it & \text{if } \lfloor\frac{m}{2}\rfloor + S_{i-1} < k < \lfloor\frac{m}{2}\rfloor + S_i. \end{cases}$$

Case B: Both l_m and w_m are even numbers

This case is demonstrated by Figure 2(a). The only difference between this case and Case A above occurs when the two covered cells at the top left corner and bottom right corner of each round of the coverage area as shown in Figure 2(a) are removed. When the servers in these two covered cells are missing in the top most uncovered round i of the coverage area, then the shortest distance between the furthest client at the top left corner of round i to the nearest covered cell (top left corner of the coverage area) is the hypotenuse of a right angled triangle whose dimensions are ia and $(i+1)a$ as shown in Table III. Using the Pythagoras theorem and after some algebra we get the result in Equation 8.

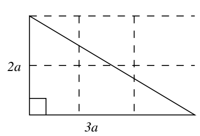


TABLE III
THE CORNER CASES FOR AN EVEN × EVEN NUMBER OF SQUARES IN THE AREA WITH $i = 2$

$$\begin{aligned} r &= a\sqrt{2i^2+2i+1} \\ &= \sqrt{2(2i^2+2i+1)}t \\ &= r_i. \end{aligned} \quad (8)$$

Hence we have

$$r = \begin{cases} \sqrt{2}t & \text{if } 0 < k \le \lfloor\frac{m}{2}\rfloor \\ 2t & \text{if } \lfloor\frac{m}{2}\rfloor < k \le \lfloor\frac{m}{2}\rfloor + S_1 - 2 \\ r_1 & \text{if } \lfloor\frac{m}{2}\rfloor + S_1 - 2 < k \le \lfloor\frac{m}{2}\rfloor + S_1 \\ 2it & \text{if } \lfloor\frac{m}{2}\rfloor + S_{i-1} < k \le \lfloor\frac{m}{2}\rfloor + S_i - 2 \\ r_i & \text{if } \lfloor\frac{m}{2}\rfloor + S_i - 2 < k < \lfloor\frac{m}{2}\rfloor + S_i. \end{cases}$$

In both of the above cases, with $k = m - m_a$, $r = \hat{r}(l, m, t, m_a)$ for some function $\hat{r}$.

The length L_p of the trajectory (path) to cover a distance of r is a function denoted as $P(r)$. Then the average time τ in seconds it takes a receiving client moving at a speed of v m/s to get its message from the nearest server is given as

$$\tau = \hat{\tau}(l, w, t, v, m_a) = \frac{P(r)}{v}. \quad (9)$$

For a random way point (RWP) mobility model, a movement of node (client) from uncovered cell to a covered cell (a cell with a server) can be assumed to be a two dimensional random walk (2DRW) [9]. For a 2DRW, a node takes $n_s = (r/\ell)^2$ steps of length ℓ to cover a distance r from uncovered cell to a covered cell [9]. Hence, using Wald's Equation [10], the expected value of the total length L_p of the path, covered by a client to get its message from the nearest server, is given by

$$L_p = \hat{L}_p(l, m, t, m_a) = \ell\frac{r^2}{\ell^2} = \frac{r^2}{\ell}. \quad (10)$$

In our experiments we consider the step size $\ell = a = \sqrt{2}t$ which is the dimension of each square cell.

The expected maximum delay τ of message delivery also depends on the the total pause time τ_p^t of the moving nodes. In the case of RWP mobility scenario, for a pause time of τ_p at every step of the path of length L_p and expected mobility step (transition) length of ℓ, the value of τ_p^t is given by

$$\tau_p^t = \frac{L_p}{\ell}\tau_p. \quad (11)$$

Once τ_p^t and the furthest length of the path L_p from the nearest server are obtained using Equations 11 and 10, the average time τ it takes a receiving client moving at a speed of v m/s to get its message from the nearest server is given as

$$\tau = \hat{\tau}(l, w, t, v, m_a) + \tau_p^t = \frac{L_p}{v} + \frac{L_p}{\ell}\tau_p. \quad (12)$$

The delay Equation 12 excludes the time it takes for the messages sent to propagate to all servers and other delay components due to congestion or interference. This is because as described using Equation 4 the propagation delay is negligible. Besides for short messaging systems the delay due to congestion is not significant. Otherwise, Equation 12 can be modified to take these delays into account.

E. Simple Reliability Model

The discussion in the above section implies that as long as $\tau \le d$, all messages with a lifetime or deadline of d sent from clients in the covered cells are successfully received. Here as described in the notations of section III-A above, d is the required message deadline which is an important quality of service (QoS) parameter. Hence, on average to satisfy the deadline d of all messages transmitted from clients at cells with server coverage using buffering of messages and mobility of receivers, we need at least $m_a = \hat{m_a}(l, w, t, v, d)$. Here $\hat{m_a}$ is an inverse of the $\hat{\tau}$ function. In the rest of this paper, we are more interested in this scenario where nodes can take advantage of buffering and mobility to receive messages sent to them.

A reliability value can also be obtained for scenarios where nodes in some cells can never get a message sent to them regardless of buffering and how fast they move. In this case, for a RWP mobility scenario, if there are κ cells whose clients can never receive messages sent to them because their $\tau > d$, then the average reliability decreases further by

$$R_d = \frac{(\kappa/m)(n/m)(m-k)}{n} = \frac{(m-k)\kappa}{m^2} = \frac{m_a\kappa}{m^2} \quad (13)$$

where κ/m is the probability that a message sent is destined in one of the κ cells with no server coverage and with $\tau > d$. More details of this case are left for future work.

Now given $m_a \geq \hat{m_a}(l, w, t, v, d)$, let's denote the distribution of the number of mobile nodes at cell (BS or server) i at time instant ω with $N_i(\omega)$ and the distribution of the rate at which client j of cell i sends packets at time ω with $R_j^i(\omega)$. The reliability R of short message delivery at time ω is given by

$$R = \hat{R}(l, w, t, m_a, d, v) = \frac{\sum_{i=1}^{m_a}\sum_{j=1}^{N_i(\omega)} R_j^i(\omega)}{\sum_{i=1}^{m}\sum_{j=1}^{N_i(\omega)} R_j^i(\omega)}. \quad (14)$$

- Assuming that the clients are uniformly distributed in the area, the average number of clients per server is $f = \frac{n}{m}$. So if at any given time one server is omitted or is disfunctional, f clients are disconnected or lost. This in turn implies that out of all the messages sent, f of them are lost on average due to lack of server nearby the sender. This is assuming that all clients send messages at the same (uniform) rate and move at a uniform (the same) speed. These uniformity assumptions can be relaxed by partitioning the node set and area into nodes with similar patterns, and then aggregating the results. Therefore, if k servers are omitted, the reliability R of message delivery is given by

$$\begin{aligned} R &= \frac{n_c}{n} = \frac{n-kf}{n} \\ R = \hat{R}(l, w, t, d, v, m_a) &= \frac{n-(m-m_a)\frac{n}{m}}{n} = \frac{m_a}{m}. \\ \Rightarrow m_a &= mR \\ &= \frac{A}{2t^2}R = \frac{lw}{2t^2}R. \end{aligned} \quad (15)$$

We call the above model a *simple model* as it is a simplified under-estimation of the actual reliability. The *simple model* under-estimates reliability as it does not take into account areas outside the square cells which are covered by each server.

By taking into account the areas outside each square covered by the server, we get a more accurate reliability model as described below. We call this model of reliability an *advanced model* of reliability.

F. The Advanced Reliability Model

In this paper we approximate the server coverage area by a square cell as shown in Table II. The same modeling approach can be used for scenarios where the server coverage area is approximated by a hexagon. For $m_a \leq \lceil m/2 \rceil$, there is no difference in reliability values between the hexagon and square cell layouts.

In the rest of the paper we consider a RWP mobility model and hence assume uniform distributions. The extensions to any other mobility models and distributions is straightforward. Hence Equation 14 can be similarly extended in the advanced reliability model for any other distribution.

In this advanced model, to derive the reliability taking into account the extra space covered by the servers outside their square areas, we first consider when half of the servers are missing, we then derive two cases where the number of missing servers is less and more than half of the total number m of servers required to give full coverage.

1) With half the number of servers needed to cover the area ($m_a = \lceil m/2 \rceil$): In this scenario, the number of remaining servers (covered squares) $m_a = \lceil m/2 \rceil$. Denote the number of missing servers with k. Also denote the total extra fraction of area which is covered by the servers outside their squares by E. The reliability R_1 for this scenario is then given by

$$\begin{aligned} R_1 &= \hat{R_1}(l, w, t, d, v, m_a) = \frac{n - k\frac{n}{m} + E\frac{n}{m}}{n} \\ &= \frac{m-k+E}{m} = \frac{m_a + E}{m}. \end{aligned} \quad (16)$$

As described in the above *simple model*, the average number of messages per unit time sent in each square under the above mentioned assumptions is $\frac{n}{m}$. This assumes that the nodes under the RWP mobility scenario are uniformly distributed. As extensively studied in the literature [2], [3], the distribution of the nodes moving following the RWP mobility scenario is not always uniform. As reported in this studies, the node distribution in the RWP mobility can be approximated by a uniform distribution under a high node speed v and high pause time τ_p^t. So in our simulation experiments, we will use a high pause time in order to approximate uniform distribution. Otherwise the average number of messages sent in a given square cell may be a certain fraction $\hat{f}$ of n.

When half of the servers needed for full coverage are missing as can be seen in Figures 2(a) to 2(c),

$$k = \lfloor \frac{m}{2} \rfloor.$$

The fraction E is obtained by adding the edge E_e, side E_s and middle E_m extra area fractions as

$$E = E_e + E_s + E_m. \quad (17)$$

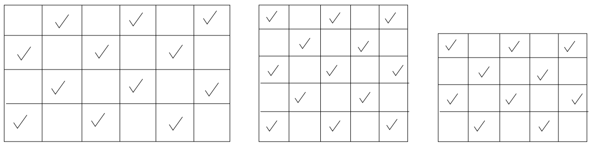
To find the values of E_e, E_s and E_m, we first derive expressions for the numbers n_e, n_s and n_m of edge, side and middle remaining servers which give extra coverage to the squares with missing servers. For a rectangular area where nodes can move, there are three distinct cases which give different expressions for these values. These cases are given as follows:

1) When the values of both l_m and w_m are even as shown in Figure 2(a). In this case the number of servers (BS)

or covered square cells n_p on the periphery of the movement area is given by

$$n_p = 2(l_m/2) + 2(w_m/2 - 1) = l_m + w_m - 2.$$

Also $n_e = 2$, $n_s = n_p - n_e$ and $n_m = \lceil m/2 \rceil - n_p$.



(a) Even by Even (b) Odd by Odd (c) Odd by Even

Fig. 2. The marked (covered) even by even, odd by odd and even by odd number of squares

2) When the values of both l_m and w_m are odd as shown in Figure 2(b). In this case the number of available servers (BS) or covered square cells n_p on the periphery of the movement area is given by

$$n_p = 2\lceil l_m/2 \rceil + 2(\lceil w_m/2 \rceil - 2) = l_m + w_m - 2.$$

Also $n_e = 4$, $n_s = n_p - n_e$ and $n_m = \lceil m/2 \rceil - n_p$.

3) When one of the values of l_m and w_m is odd and the other is even as shown in Figure 2(c). In this case the number of servers (BS) or covered square cells n_p on the periphery of the movement area is given by

$$n_p = 2\lceil l_m/2 \rceil - 1 + 2(\lceil w_m/2 \rceil - 1) = l_m + w_m - 2.$$

Also $n_e = 4$, $n_s = n_p - n_e$ and $n_m = \lceil m/2 \rceil - n_p$.

We next derive the fractions E_e, E_s and E_m of areas covered by each BS outside its square. As shown in Figure 3 each BS in the middle of the movement area covers 4 sectors from the other squares whose BS is missing. Each BS at each edge square covers 2 sectors from the squares of the missing BS and each BS on the sides of the movement area covers 3 sectors. The area of each square in the rectangular movement

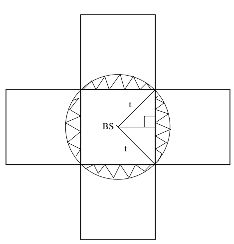


Fig. 3. The sectors covered by a middle server (BS)

area is given by $A_{sqr} = a^2 = 2t^2$ as shown in Equation 2. The area of each sector as shown in Figure 3 is given by $A_{sct} = \pi t^2/4 - t^2/2$. Hence, the fraction of nodes moving and exchanging messages in each sector of each square is

$$\phi = \frac{A_{sct}}{A_{sqr}} = \frac{\pi t^2/4 - t^2/2}{2t^2} = \frac{\pi/2 - 1}{4}. \quad (18)$$

Hence as discussed above, we have the following equations.

$$\begin{aligned} E_e &= n_e \times 2 \times \phi & (19)\\ E_s &= n_s \times 3 \times \phi & (20)\\ E_m &= n_m \times 4 \times \phi. & (21) \end{aligned}$$

We next present the reliability for the case where less than half of the number of servers needed to cover the rectangular area.

2) With less than half of the number of servers needed to cover the area ($m_a < \lceil m/2 \rceil$): In this case the number m_a of remaining BS is less than half the total number $\lceil m/2 \rceil$ of BS required to give full coverage. Hence, we have the following three cases.

Edge Case: $0 < \lceil m/2 \rceil - m_a \leq n_e$
In this case there will be $\lceil m/2 \rceil - m_a$ more BS missing from the movement area. These first missing servers are chosen to be the once at the edges. One justification for this is that a BS at a square edge covers only 2 sectors of the other uncovered squares around it. And this is the smallest number of sectors of other square areas a BS covers when compared with the side and middle covered cells. The reliability of this case decreases from when there are about half BS proportionally. Hence the reliability R_2 of this case is given as the reliability R_1 of the case $m_a = \lceil m/2 \rceil$ with half of the BS missing minus the reliability loss $R_L^e(l, w, t, d, v, m_a)$ due to the additional $\lceil m/2 \rceil - m_a$ missing edge BS. This reliability loss is due to the uncovered edge squares and the sectors of the neighboring cells which were supposed to be covered by the BS at these squares. Hence we have

$$\begin{aligned} R_L^e(l, w, t, d, v, m_a) &= \frac{(\lceil m/2 \rceil - m_a)(1 + 2\phi)n/m}{n} \\ &= \frac{(\lceil m/2 \rceil - m_a)(1 + 2\phi)}{m} \quad (22) \end{aligned}$$

and the reliability R_2 of this case with more than half of the BS missing is given as

$$R_2 = \hat{R}_2(l, w, t, d, v, m_a) = R_1 - R_L^e(l, w, t, d, v, m_a) \quad (23)$$

where R_1 is given by Equation 16.

Side Case: $n_e < \lceil m/2 \rceil - m_a \leq n_p$
In this case more than the edge BS are missing. Hence the reliability R_2 of the above case is further reduced by the reliability loss R_L^s due to the additional missing side BS to get the new reliability R_3 of this case. Hence we have

$$\begin{aligned} R_L^s(l, w, t, d, v, m_a) &= \frac{(\lceil m/2 \rceil - m_a - n_e)(1 + 3\phi)n/m}{n} \\ &= \frac{(\lceil m/2 \rceil - m_a - n_e)(1 + 3\phi)}{m} \quad (24) \end{aligned}$$

and the reliability R_3 of this case with less than half of the BS missing with more than edge BS losses is given as

$$\begin{aligned} R_3 &= \hat{R}_3(l, w, t, d, v, m_a) \\ &= \hat{R}_2(l, w, t, d, v, \lceil m/2 \rceil - n_e) - R_L^s(l, w, t, d, v, m_a). \end{aligned}$$

Middle Case: $n_p < \lceil m/2 \rceil - m_a \leq \lceil m/2 \rceil$
In this case we have even more BS missing from the movement

area. In addition to more than half of the missing BSs and the missing periphery BSs, we have $\lceil m/2 \rceil - m_a - n_p$ middle BSs missing resulting in an additional $R_L^m(l, w, t, d, v, m_a)$ reliability loss. Hence we have

$$R_L^m(l, w, t, d, v, m_a) = \frac{(\lceil m/2 \rceil - m_a - n_p)(1 + 4\phi)n/m}{n} = \frac{(\lceil m/2 \rceil - m_a - 2)(1 + 4\phi)}{m} \quad (25)$$

and the reliability R_4 of this case with less than half of the BS missing with more than periphery BS losses is given as

$$\begin{aligned} R_4 &= \hat{R}_4(l, w, t, d, v, m_a) \\ &= \hat{R}_3(l, w, t, d, v, \lceil m/2 \rceil - n_p) - R_L^m(l, w, t, d, v, m_a). \end{aligned}$$

3) With more than half of the number of servers needed to cover the area ($m_a > \lceil m/2 \rceil$) : With more than half of the total number m of the BSs needed to cover rectangular movement area, we have an additional gain of reliability. With each additional ($m_a > \lceil m/2 \rceil$) BS covering each square area we need to exclude the sectors of each newly covered square which were previously covered by the neighboring BS in the case when $m_a = \lceil m/2 \rceil$. This is to avoid double-counting of the newly covered square areas.

We next present expressions for the total numbers n_e^r, n_s^r and n_m^r of remaining (uncovered) edge, side and middle squares to be covered by the additional BSs. There are 4 corner squares in a rectangle. If n_e of them are covered by half of the BS needed to provide full coverage, then $n_e^r = 4 - n_e$. Similarly $n_p^r = 2l_m + 2w_m - 4 - n_p = 2(l_m + w_m - 2) - n_p$ and $n_m^r = \lfloor m/2 \rfloor - n_p^r$ where $n_s^r = n_p^r - n_e^r$.

Edge Case: $0 < m_a - \lceil m/2 \rceil \leq n_e^r$
In this case, we have an additional $m_a - \lceil m/2 \rceil$ covered squares and 2 covered sectors with each square which have to be excluded to avoid double-counting. The placement of the additional BSs starts at the square corners as they are the areas with less sectors covered. For instance if we take the middle cell missing, four sectors from its square area may be covered by the neighboring servers (BS). However only two sectors from the edge cell area may be covered by the two neighboring servers. Hence the reliability gain $R_G^e(l, w, t, d, v, m_a)$ is given as

$$R_G^e(l, w, t, d, v, m_a) = \frac{(m_a - \lceil m/2 \rceil)(1 - 2\phi)n/m}{n} = \frac{(m_a - \lceil m/2 \rceil)(1 - 2\phi)}{m} \quad (26)$$

and the reliability R_5 of this case with only less than half of the BS missing is given as

$$R_5 = \hat{R}_5(l, w, t, d, v, m_a) = R_1 + R_G^e(l, w, t, d, v, m_a). \quad (27)$$

Side Case: $n_e^r < m_a - \lceil m/2 \rceil \leq n_p^r$
In this case, we have an additional $m_a - \lceil m/2 \rceil - n_e^r$ covered side squares and 3 covered sectors with each square. Hence the reliability gain $R_G^s(l, w, t, d, v, m_a)$ is given as

$$R_G^s(l, w, t, d, v, m_a) = \frac{(m_a - \lceil m/2 \rceil - n_e^r)(1 - 3\phi)n/m}{n} = \frac{(m_a - \lceil m/2 \rceil - n_e^r)(1 - 3\phi)}{m} \quad (28)$$

and the reliability R_6 of this case with only less than half of the BS missing is given as

$$\begin{aligned} R_6 &= \hat{R}_6(l, w, t, d, v, m_a) \\ &= \hat{R}_5(l, w, t, d, v, \lceil m/2 \rceil + n_e^r) + R_G^s(l, w, t, d, v, m_a). \end{aligned}$$

Middle Case: $n_p^r < \lceil m_a - m/2 \rceil \leq n_p^r + n_m^r$
In this case, we have an additional $m_a - \lceil m/2 \rceil - n_p^r$ covered middle squares and 4 covered sectors with each square. Hence the reliability gain $R_G^s(l, w, t, d, v, m_a)$ is given as

$$R_G^m(l, w, t, d, v, m_a) = \frac{(m_a - \lceil m/2 \rceil - n_p^r)(1 - 4\phi)n/m}{n} = \frac{(m_a - \lceil m/2 \rceil - n_p^r)(1 - 4\phi)}{m} \quad (29)$$

and the reliability R_6 of this case with only less than half of the BS missing is given as

$$\begin{aligned} R_7 &= \hat{R}_7(l, w, t, d, v, m_a) \\ &= \hat{R}_6(l, w, t, d, v, \lceil m/2 \rceil + n_p^r) + R_G^m(l, w, t, d, v, m_a). \end{aligned}$$

IV. Simulation Validation

To validate our analytical models of short message transfer reliability we have modified the NS2 simulation package to simulate senders and receivers of messages along with servers which control the message transfers. In the experiment presented in this paper we have used a star topology similar to the one shown in sample Figure 4 with one server at the center and all other servers connecting to it. In this figure, S represents sender node, R represents receiver node and C represents session controller (server) nodes.

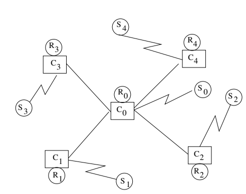


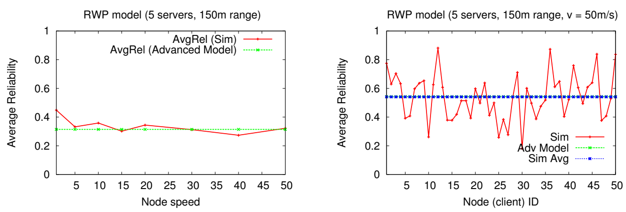
Fig. 4. Sample network topology

In this simulation setup we consider a $1061mx1061m$ rectangular area with $l = m = 1061m$, where the nodes can move. We use 25 to 50 senders and 25 to 50 receivers. The message lifetime (deadline) d ranges from 60s to 120s and beacon interval is 0.1s. We use a random way point mobility model with the average speed v of nodes ranging from 1m/s to 50m/s. We use a higher pause time τ_p^t of up to to 90s for the case where receivers are mobile to emulate a uniformly distribution of the nodes in the mobility area. The transmission range t for our experiments is 150m. The results in this section are for a fixed packet (chat message) size of $24+$IPHDRLEN

Bytes, where $20 \leq$ IPHDRLEN ≤ 60. The chat message is generated at a uniform rate of $1packet/5s$. We next present the numerical results for cases when the number m_a of servers is 5 and 9.

A. With 5 servers

For the first case with 5 servers, Figures 5(a) and 5(b) show that the *Advanced Reliability Model* gives accurate estimation of the average reliability when 50 senders are mobile and 50 receivers are static. Figure 5(b) shows that the average reliability of the advanced analytical model (Adv Model) coincides with the simulation average (Sim Avg).

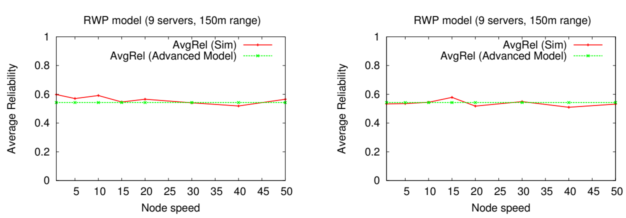


(a) With variable speed (b) With a speed of 50m/s

Fig. 5. Our Model vs simulation results

B. With 9 servers

Figure 6(a) shows mobility results when 50 senders are mobile and 50 receivers are static. Figure 6(b) shows results where 25 senders and 25 receivers are mobile. Both experiments consider 9 servers and show that our analytical model gives accurate estimation of the average reliability of message delivery.



(a) With variable sender speed (b) With variable sender and receiver speeds

Fig. 6. Our Model vs simulation results

Figure 7 shows the case of 9 servers with 50 senders and 50 receivers mobile at a speed of 30m/s. From the plot it can be seen that the average reliability using the analytical model (Adv Model) almost coincides with the average of the simulation (Sim Avg).

In all our experiments the number of servers m_a was such that maximum message deadline d didn't exceed the average maximum lifetime τ_p^t as calculated with Equation 12. This allowed moving receivers to eventually receive the messages sent to them before the deadline expires. This explains why the average reliabilities presented in Figures 6(a) and 6(b) for cases where receivers are mobile and static are almost similar. As discussed in Section III-D, for cases where the message deadline exceeds, Equation 13 can be used to account for the loss (reduction) in reliability.

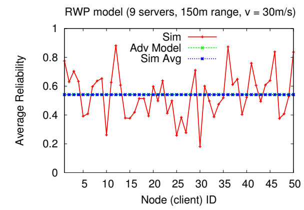


Fig. 7. Our Models vs simulation results with a sender and receiver speeds of 30m/s

V. Summary and Future Work

In this paper we present closed form expressions for the reliability of message delivery from one mobile node to another mobile node as multivariate function of the dimensions of the rectangular movement area, client speed, message lifetime, the number of servers and their transmission range. Simulation results show that our analytical model gives accurate estimation of the average message reliability.

Making the simple extension of our work to cases where message life time expires and when the cell shapes are hexagon is left for future work. We also plan to validate our analytical model of the reliability with real mobile network experiments using Android phones.

References

[1] Bettstetter, C., Hartenstein, H., and Pérez-Costa, X. Stochastic properties of the random waypoint mobility model. *Wirel. Netw. 10* (September 2004), 555–567.

[2] Bettstetter, C., Resta, G., and Santi, P. The node distribution of the random waypoint mobility model for wireless ad hoc networks. *IEEE Transactions on Mobile Computing 2* (July 2003), 257–269.

[3] Blough, D. M., Resta, G., and Santi, P. A statistical analysis of the long-run node spatial distribution in mobile ad hoc networks. *Wirel. Netw. 10* (September 2004), 543–554.

[4] Cheng, A., Evans, M., and Singh, H. Inside Twitter: An In-Depth Look Inside the Twitter World. *Sysomos Inc., A Marketwire Company* (June 2009).

[5] Pongthawornkamol, T. *Reliability and timeliness analysis of content-based publish/subscribe systems*. PhD thesis, University of Illinois at Urbana-Champaign, Urbana, Illinois, 2010.

[6] Pongthawornkamol, T., Nahrstedt, K., and Wang, G. Probabilistic QoS modeling for reliability/timeliness prediction in distributed content-based publish/subscribe systems over best-effort networks. In *ICAC '10* (New York, NY, USA, 2010), ACM, pp. 185–194.

[7] Saint-Andre, P. Extensible Messaging and Presence Protocol (XMPP): Instant Messaging and Presence (draft-ietf-xmpp-3921bis-12). *XMPP Standards Foundation*, 12 (2010).

[8] Tseng, Y.-C., Lin, T.-Y., Liu, Y.-K., and Lin, B.-R. Event-driven messaging services over integrated cellular and wireless sensor networks: prototyping experiences of a visitor system. *IEEE Journal on Selected Areas in Communications 23*, 6 (2005), 1133–1145.

[9] Weisstein, E. W. Random Walk–2-Dimensional. *MathWorld–A Wolfram Web Resource*.

[10] Weisstein, E. W. Wald's Equation. *MathWorld–A Wolfram Web Resource*.

[11] Xiaomin, M., Yun, L., and Kishor, T. S. Design and Performance Analysis of a New Soft Handoff Scheme for CDMA Cellular Systems. *IEEE TRAN. ON VEHICULAR TECH. 55*, 5 (2006), 1603–1612.

[12] Zerfos, P., Meng, X., Wong, S. H., Samanta, V., and Lu, S. A study of the short message service of a nationwide cellular network. In *IMC '06* (New York, NY, USA, 2006), ACM, pp. 263–268.