A Study of Spatial Packet Loss Correlation in 802.11 Wireless Networks

Zhe Wang, Mahbub Hassan
School of Computer Science and Engineering
University of New South Wales, Sydney 2052, Australia
Email: {zhewang, mahbub}@cse.unsw.edu.au

Tim Moors
School of Electrical Engineering and Telecommunications
University of New South Wales, Sydney 2052, Australia
Email: t.moors@unsw.edu.au

Abstract—This paper examines the spatial correlation of packet loss events in IEEE 802.11 wireless networks for broadcast communications. We discuss limitations of previously used metrics to measure spatial loss correlation and show that the entropy correlation coefficient, which is based on the mutual information concept of Information Theory, can overcome these limitations. In our experiments, we find that the packet losses among closely located receivers are highly correlated when signal strength is high, but the correlation decreases with decreasing signal strength. The losses become totally independent once the signal strength falls below a threshold. As a first step to model the relationship between spatial loss correlation and signal strength, we find that the relationship can be approximated as a Gaussian cumulative distribution function.

I. INTRODUCTION

Packet loss is an important issue for IEEE 802.11 wireless networks. A great amount of research has been done to study the average packet loss rate, or the packet loss probability. However, a more fine-grained study of the issue of packet loss correlation also attracts the interest of the many researchers [1]–[3].

There are two kinds of packet loss correlation, namely temporal correlation and spatial correlation. The temporal loss correlation measures the packet loss pattern over time. It is often related to the issue of consecutive packet losses or the burstiness of packet loss by a single receiver. However, for multicast/broadcast communication, in which a single packet has multiple receivers, it is also useful to measure the correlation of losses by different receivers. We focus on such spatial loss correlation between receivers of a single transmission.

The spatial correlation between receivers is important for its effect on the performance of error recovery protocols. If there is low correlation in loss between receivers that are near each other, then when one receiver loses a packet, it may be able to recover the lost packet from a nearby receiver. Thus, local retransmission is more effective when losses are less spatially correlated. The extent to which losses among multiple receivers are disjoint also influences error recovery when network coding is used to combine (e.g., XOR) multiple lost packets into a single retransmission [4], [5].

Despite the importance of spatial loss correlation in wireless networks, not much is known about this issue. In open literature, the 802.11 spatial loss correlation has been reported in [2] and [3]. Reis et al [2] stated qualitatively that the loss between receivers is “roughly independent” but did not present quantitative data about the loss correlation. Salyers et al [3] presented the results of a small range (distances of 3m∼12m and packet loss rate of 2%~10%) experiment, and found, using conditional loss probability as a measure of loss correlation, that the loss correlation increases with distance. However, as we will demonstrate later in this paper, conditional loss probability alone cannot truly capture the statistical dependence among the losses observed at different receivers, hence may invalidate some of the conclusions reported in [3].

Clearly, further studies are needed to explore the issue of spatial packet loss correlation in wireless networks. Given that in many applications, e.g., 802.11-based vehicular safety communication, the packet loss rate is up to 40% due to distance between the packet sender and receivers can be hundreds of meters ( [6], [7]), it is important to characterise spatial loss correlation for a wider range of distances than reported in [3]. In this paper, we study the effect of distance on packet loss correlation by a series of experiments.

We make the following key contributions:

- We discuss the limitations of the metrics used to measure loss correlation in previous works and propose a new metric to overcome these limitations.
- We conduct a series of experiments representing a wide range of distances between the transmitter and the receivers.
- Applying the proposed metric, we find that the spatial packet loss correlation in our 802.11 network experiments decreases with the distance between sender and receivers.
- We take a first step to analytically model the loss correlation. We find that with signal strength as the independent variable, the proposed loss correlation metric can be approximated as a Gaussian cumulative distribution function.

The rest of this paper is organized as follows: We review the related work in Section II. The limitation of previously used metrics to measure spatial loss correlation are discussed and the proposed metric is presented in Section III. We describe our experiments and discuss the findings in Section IV. We conclude in Section V.

1In the rest of this paper, the term loss correlation specifically refers to the spatial packet loss correlation.
II. RELATED WORK

Miu et al. [1] have investigated whether separate 802.11 transmitters have correlated packet losses when observed at a single receiver. The metric used in [1] to measure the correlation is based on the conditional and unconditional loss probability. If $A$ is the event of the receiver losing a packet from transmitter 1 and $B$ is the event of the receiver losing a packet from transmitter 2, then the level of correlation can be measured by comparing the probability $P(A)$ and $P(A|B)$. If $P(A)$ equals $P(A|B)$, then the packet loss of is totally uncorrelated (probabilistic independent). At the other extreme, if $P(A|B)$ equals one, then the packet loss is totally correlated. The metric used to measure the loss correlation is the difference between $P(A|B)$ and $P(A)$. The level of correlation is decided by the position of $P(A|B)$ in the range $[P(A), 1]$. The further $P(A|B)$ is from $P(A)$ and closer to 1, the higher the correlation is. Although the loss correlation issue we are investigating is different from [1], the metric used in this work can also be applied to the case of one transmitter and two receivers. In our case, for any packet from the transmitter, $A$ is the event of receiver 1 losing the packet and $B$ is the event of receiver 2 losing the packet.

Reis et al [2] observed that the spatial loss correlation of 802.11 network in a office building is low. However, neither the metric used to measure the correlation nor the loss correlation data is shown in the paper.

The most related work to our study is [3], in which the authors study the loss correlation for different 802.11 wireless cards at different distances between the sender and receivers. They measure the correlation using the conditional loss probability alone, instead of using the difference between the conditional and unconditional loss probabilities as in [1]. Using this metric, the authors report that the loss correlation increases when distance increased from 3m to 12m.

In summary, previous works have measured the 802.11 packet loss correlation using different metrics. Before we start our study on how the loss correlation changes with the distance between sender and receivers, it is necessary to choose an appropriate metric to numerically measure the loss correlation.

III. METRIC FOR SPATIAL LOSS CORRELATION

In this section, we propose a metric to measure the loss correlation. We start by discussing the properties that the metric should have.

The first property that the loss correlation metric should have is that it should indicate the statistical dependence of the packet loss between two receivers. In probability theory, the definition of two events $A$ and $B$ being independent is: $P(AB) = P(A)P(B)$ [8]. Using the conditional probability $P(A|B)$ alone to measure the loss correlation as in [3] is problematic because only when $P(A|B) = P(A)$, i.e., $\frac{P(AB)}{P(B)} = P(A)$, can we get $P(AB) = P(A)P(B)$. The conditional probability $P(A|B)$ does not indicate whether the packet losses of the two receivers are independent without considering its relationship with unconditional probability $P(A)$. The metric used in [1], which is the relationship between conditional loss probability $P(A|B)$ and unconditional loss probability $P(A)$, satisfies this property. On one hand, if the ratio $\frac{P(AB)}{P(A)} = 1$, we can get $P(AB) = P(A)P(B)$, which means the packet loss is independent (uncorrelated). On the other hand, if the ratio $\frac{P(AB)}{P(A)} = \frac{1}{P(A)}$, we can get $P(AB) = 1$, which means $A \subseteq B$, i.e., the packet loss is 100% correlated. The ratio of the conditional probability to the unconditional probability, $\frac{P(AB)}{P(A)}$, is bounded by $[1, \frac{1}{P(A)}]$. The higher this ratio, the higher the packet loss is correlated between the two receivers.

Instead of using the conditional loss probability, [9] uses the Pearson correlation coefficient to measure the spatial loss correlation in wired multicast networks. The Pearson correlation coefficient is a measure of linear dependence between two random variables [8], which ranges from -1 to 1. If the coefficient is 1 or -1, the two random variables have perfect linear positive or negative relationship (100% correlated). If the coefficient is 0, the two random variables have no linear dependence. The coefficient being 0, i.e., no linear dependence, does not guarantee the statistical independence. In some cases, the two random variables ($X$ and $Y$) are dependent, e.g., $Y = X^2$ while the coefficient is 0. Therefore, the Pearson correlation coefficient is inadequate to measure the statistical dependence of the packet loss between two receivers.

The second property that the loss correlation metric should have is that given the packet loss events ($A$ and $B$) of two receivers, the metric ($\rho$) should have a unique value no matter which receiver is chosen as $A$ or $B$, i.e., $\rho_{AB} = \rho_{BA}$. The ratio of conditional loss probability over unconditional loss probability does not satisfy this property when $\frac{P(AB)}{P(A)} \neq \frac{P(B)}{P(A)}$. By using conditional/unconditional loss probability ratio as the metric, we might have two possible measures for the loss correlation between the same pair of receivers.

The third property the loss correlation metric should have is that the metric should have fixed upper and lower bounds. Without this property, it is hard to compare the level of loss correlation between two different pairs of receivers. For example, assume that we have two pairs of receivers ($R_1$, $R_2$), and ($R_1'$, $R_2'$). For ($R_1$, $R_2$), the packet loss events are $A$ and $B$ respectively, and for ($R_1'$, $R_2'$), the packet loss events are $A'$ and $B'$ respectively. By using the ratio of unconditional/conditional loss probabilities as the metric, we cannot compare the level of loss correlation of these two pairs if $P(A) \neq P(A')$ because the ratios: $\frac{P(AB)}{P(A)}$ and $\frac{P(A'B')}{P(A')}$, have different upper bounds ($\frac{1}{P(A)}$ and $\frac{1}{P(A')}$). Suppose that the packet loss of both these two pairs are 100% correlated (they have the same correlation level), comparing $\frac{P(AB)}{P(A)}$ (equals $\frac{1}{P(A)}$ in this case) and $\frac{P(A'B')}{P(A')}$ (equals $\frac{1}{P(A')}$. In this case) will lead to the wrong conclusion that the loss correlation of these two pairs are different.

The metrics used in previous works do not have all the three properties discussed above. In the following, we propose a new metric that exhibits all these three properties.

The entropy correlation coefficient [10] is a metric to
measure the dependency of two random variables, which is based on the entropy concept of information theory. If we have two random variables $X$ and $Y$, the entropy correlation coefficient ($\rho_{XY}$) is defined as:

$$\rho_{XY} = \frac{2I_{XY}}{H_X + H_Y}$$

(1)

where $H_X$ and $H_Y$ are the entropies of $X$ and $Y$ respectively, and $I_{XY}$ is the mutual information of $X$ and $Y$. If we define the random variables $X$ and $Y$ as: $X = 0$ if a packet is received by receiver $R_1$, $X = 1$ if the packet is lost by $R_1$; $Y = 0$ if the packet is received by receiver $R_2$, $Y = 1$ if the packet is lost by $R_2$, then we have:

$$H_X = - \sum_{x=0,1} P(X = x) \log_2 P(X = x)$$

$$H_Y = - \sum_{y=0,1} P(Y = y) \log_2 P(Y = y)$$

$$I_{XY} = \sum_{x=0}^{1} \sum_{y=0}^{1} P(X = x, Y = y) \log_2 \frac{P(X = x, Y = y)}{P(X = x)P(Y = y)}$$

The entropy correlation coefficient ($\rho_{XY}$) in (1) is the measure of dependency of the packet loss of $R_1$ and $R_2$ and it has the following properties:

- $\rho_{XY} = \rho_{YX}$, it has the same value no matter which receiver is chosen as $X$ and which is chosen as $Y$.
- $0 \leq \rho_{XY} \leq 1$, the upper and lower bounds are constants.
- $\rho_{XY} = 1$ if and only if $X$ and $Y$ are completely dependent, i.e., if we know the value of $X$, we can predict the value of $Y$ with certainty.
- $\rho_{XY} = 0$ if and only if $X$ and $Y$ are independent, i.e., knowing the value of $X$ does not help at all to predict the value of $Y$.

In the following section, we conduct a series of experiments to study the loss correlation between two closely located receivers. By measuring the loss correlation using different metrics, we show that the entropy correlation coefficient is a better metric than the ones used in previous works.
the receivers, we repeat the experiment with different transmission power levels. Starting from the maximum transmission power (17 dBm) allowed by the wireless card, the transmission power is reduced by 1 dBm for each experiment until 5 dBm. For each transmission power, the transmitting node broadcasts 20,000 packets with 20 ms gap between two consecutive packets. In order to observe a large range of packet loss rates, the external antenna is not mounted on the transmitting node (as shown in Fig. 1). The receiving nodes are set to work in monitor mode so that the Receiving Signal Strength Indicator (RSSI) is included in the captured packet traces.

Although the two receivers are very close to each other, which makes the path loss and channel fading have the same effect on them, they have different loss rates. By analyzing the RSSI of received packets as shown in Fig. 2, we find that the two receivers have slightly different signal to noise ratios (In Madwifi, the reported RSSI for each packet is equivalent to the signal to noise ratio [14]). The average RSSI of receiver 1 is always lower than that of receiver 2. The difference remains the same even when we swap the locations and the external antennas. This might be caused by the variation of analog part of the hardware.

B. Conditional/Unconditional Loss Probability

The conditional and unconditional packet loss probabilities at different transmission powers are shown in Fig. 3. Note that, the transmission power range we applied increased unconditional loss probability by about 1000 folds from 0.001 at 17 dBm to 0.997 at 5 dBm. This would reflect a large variation in the transmitter-receiver distance for a fixed transmission power (In [3], the unconditional probability changes from 0.02 to 0.10, an increase of about 5 fold, when the distance increases from 3m to 12m).

While the unconditional loss probabilities monotonically decrease with transmission power, the conditional loss probabilities are “V” shape curves. When the transmission power is low, the conditional loss probability alone cannot capture the statistical dependence among the losses of the two receivers. For example, when the transmission power is 8dBm, $P(A)=0.93$ and $P(A|B)=0.93$; when the transmission power is 14dBm, $P(A)=0.009$ and $P(A|B)=0.93$. Although the conditional loss probabilities are the same, the statistical dependence of packet loss of these two cases is fairly different. For 8dBm, we have $P(A|B)=P(A)$, suggesting that there is little dependency, whereas for the case of 14dBm, $P(A|B)$ is much larger than $P(A)$, suggesting that the packet losses at two receivers are highly dependent.

Fig. 3: The conditional and unconditional loss probabilities of the two receivers. $P(A)$ and $P(B)$ are the unconditional loss probabilities of receiver 1 and 2 respectively. $P(A|B)$ and $P(B|A)$ are the conditional probabilities of a receiver losing a packet given that the other receiver loses the packet.

Fig. 4: The ratio of conditional/unconditional loss probability (in log-scale).
C. Entropy Correlation Coefficient

From the packet trace collected by the two receivers, we derive the joint distribution of $X$ and $Y$ defined in Section III and calculate the entropy correlation coefficient using (1). The results are shown in Fig. 5 (along with the packet reception probabilities at the receivers). We can see that the loss correlation decreases monotonically with decreasing transmission power. When the transmission power is high, i.e., above 16dBm, the entropy correlation coefficient is one, which means that the two receivers lose the same set of packets and if one receiver loses (or receives) a packet, we know the other also loses (or receives) it. Loss correlation starts to decrease as transmission power is reduced. When the transmission power is less than 8dBm, the entropy correlation coefficient is almost zero, which means that the two receivers independently lose packets and the knowledge of one receiver losing (receiving) a packet is not useful to predict the loss (reception) at the other receiver.

Another interesting observation we make from Fig. 5 is that the curves of both the entropy correlation coefficient and the packet reception probabilities have the “S” shape, albeit the centre points and steepness are different. This observation suggests that we may be able to model loss correlation in a similar fashion the packet reception probability is modelled. We first discuss the possible reasons for the packet reception probabilities to exhibit the “S” shape.

It is known that, if the signal strength attenuates according to log-normal shadowing model, then the “S” shaped Gaussian cumulative distribution function (CDF) can be used to determine the probability that the received signal strength (RSS) is greater than a threshold $R(R_{th})$ [15]. If our experimental radio environment is characterised by log-normal shadowing model, and if we assume that packets are received correctly only when the RSS exceeds a threshold (this assumption is used widely in the literature and simulation tools [16]), the packet reception probability can be approximated by Gaussian CDF. To investigated the radio model of our experimental environment, we plot the RSSI distribution for different transmission power in Fig. 6. Log-normal shadowing implies that RSS at specific transmitter-receiver separation is Gaussian (normal) distributed with the distance dependent mean and a constant standard
probability from the experiment to the function steep).

The position of the center point of the curve (increasing plotted in Fig. 7. It can be seen that parameter 

According to the log-normal shadowing model [15], the received signal strength (R) can be expressed as:

$$ R = T - L + N $$

where $T$ is the transmission signal strength, $L$ is the path loss and $N$ is a Gaussian random variable with zero mean and standard deviation $\sigma$ ($T$, $L$ and $\sigma$ are measured in dBm). Given that $T$ and $L$ are static variables, $R$ is Gaussian distributed with mean $\mu_0 = T - L$ and standard deviation $\sigma$.

Based on (2), if we assume that there is a threshold of $R(R_{th})$ above which the packet will be successfully received, the packet reception probability ($r$) can be derived as:

$$ r = P(R > R_{th}) = P(N > R_{th} - T + L) = P(N < T - R_{th} - L) $$

Since $N$ is Gaussian distributed with mean 0 and standard deviation $\sigma$, we get:

$$ r = \frac{1}{2}(1 + \text{erf}(\frac{T - \mu}{\sqrt{2}\sigma})) $$

(3)

where

$$ \mu = R_{th} + L $$

$$ \text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp(-x^2) dx. $$

Equation (3) has the same form as the CDF ($\Phi$-function) of a Gaussian random variable with mean $\mu$ and standard deviation $\sigma$:

$$ \Phi_{\mu,\sigma}(X) = \frac{1}{2}(1 + \text{erf}(\frac{X - \mu}{\sqrt{2}\sigma})) $$

(4)

The $\Phi_{\mu,\sigma}(X)$ function with different parameters values are plotted in Fig. 7. It can be seen that parameter $\mu$ controls the position of the center point of the curve (increasing $\mu$ will shift the center point rightwards). Parameter $\sigma$ controls the steepness of the curve (increasing $\sigma$ will make the curve less steep).

Using the curve fitting tool [19], we fit the packet reception probability from the experiment to the function $\Phi_{\mu,\sigma}(X)$ with unknown parameters $\mu$ and $\sigma$. The curve fitting results in Fig. 8 shows that the packet loss rate at different transmission powers can be accurately modeled by $\Phi_{\mu,\sigma}(X)$ function with parameters $\mu$ and $\sigma$ being slightly different for different receivers.

Now that we have reasonable clues why packet reception probability follows a Gaussian CDF, we proceed to examine the curve of entropy correlation coefficient. Again using the curve fitting tool [19], we fit the entropy correlation coefficient to function $\Phi_{\mu,\sigma}(X)$. Fig. 9 shows that, like the packet reception probability, the entropy correlation coefficient also has a good fit to Gaussian CDF except the fact that the values of both its parameters, $\mu$ and $\sigma$, are larger than those of the packet reception probability. How to derive the parameters $\mu$ and $\sigma$ of the Gaussian CDF that approximates spatial loss correlation remains a topic of ongoing research.

V. CONCLUSION AND FUTURE WORKS

Using indoor experiments, we have studied the spatial correlation of packet loss in IEEE 802.11 wireless networks when a packet transmitted by a single transmitter is to be received by multiple receivers. We have found that the metric used to measure the spatial loss correlation is an important factor. Specifically, we have shown that simply using conditional loss probabilities to measure loss correlation can lead to misleading conclusions. We have proposed and demonstrated that the entropy correlation coefficient, which is based on the mutual information concept of Information Theory, is a good metric to measure spatial packet loss correlation in wireless networks. Using this metric, we have discovered that the received signal strength (distance between the transmitter and the receivers) is a key factor that affects the spatial loss correlation. In our experiments, the packet losses among closely located receivers are highly correlated when the signal is strong, but the correlation falls with decreasing signal strength. The losses become totally independent once the signal strength falls below a threshold. Our findings suggest that the decay of entropy correlation coefficient can be approximated as a Gaussian CDF, the same function that models the packet reception probability at the receivers in a log-normal shadowing radio propagation environment. This finding serves as a first step to analytically model the spatial packet loss correlation.

In future work, it will be useful to conduct more experiments to investigate the impact of 802.11 mode (a/b/g), data rate, packet size, and the distance between the two receivers, on the
spatial packet loss correlation. These experiments may provide useful insight into the factors that affect the key features of the Gaussian CDF, i.e., the centre point and the steepness of the curve. In addition to studying the loss correlation in the “interference free” scenario, it may be worth considering the case when packet losses are also caused by collisions, which is common for applications where every node is a potential broadcaster (such as vehicular safety communication).

**REFERENCES**


