Clustering Coefficient of Wireless Ad Hoc Networks and the Quantity of Hidden Terminals

Li-Hsing Yen, *Member, IEEE,* and Yang-Min Cheng

***Abstract*— Clustering coefficient has been proposed to characterize complex networks. Hidden terminals may degrade the performance of CSMA (carrier sense multiple access) protocol. This letter computes analytically the clustering coefficient and the quantity of hidden terminals for ad hoc networks. The former turns out to be a constant and the latter is proportional to n^3p^2, where n is the number of nodes and p is the link probability. The connection between them has been established, and simulation results confirm our analytic work.**

I. Introduction

NETWORKS of complex topology such as social networks and the Internet were traditionally modeled as random graphs [1]. In Watts and Strogatz's pioneer work [2], they recognized that many real systems are better described as 'small-world' networks rather than random graphs. Small-world networks differ from random graphs in the tendency of clustering, or cliqueness, which is the extent to which a node's neighbors are also neighbors to each other. Specifically, for node i having $m_i \geq 2$ neighbors, at most $C(m_i, 2)$ links may exist between these neighbors. Let E_i be the total number of links that exist among i's neighbors. Node i's clustering coefficient, c_i, is defined to be $E_i/C(m_i, 2)$. The clustering coefficient of the whole network is the average of all individual c_i's.

Clustering coefficients of random graph, regular network [2], and small-world network have been well investigated [3]. To the best knowledge of the author, however, the clustering coefficient of mobile ad hoc (multi-hop) networks (MANETs) has not yet been known. In this letter, we have computed analytically the clustering coefficient of MANET under the assumption of uniform location model (Section II).

Hidden terminals refer to a pair of nodes that cannot sense each other but have at least one common neighbor node [4]. Transmission collisions may occur between hidden terminals, which cannot be prevented by carrier sensing. The existence of hidden terminals thus degrades the performance of CSMA (carrier sense multiple access) protocol substantially [5]. There have been extensive schemes proposed to deal with hidden terminal problems (e.g., RTS/CTS-like handshake [6], [7]). However, little research has been done on quantifying hidden terminals for a given MANET. We also have analyzed the number of hidden terminals and found its connection to the clustering coefficient (Section III).

Manuscript received July 20, 2004. This work has been supported by the National Science Council, ROC, under grant NSC 93-2213-E-216-024.

The authors are with the Department of Computer Science and Information Engineering, Chung Hua University, Hsinchu, Taiwan, ROC (e-mail: lhyen@chu.edu.tw).

II. Clustering Coefficient of MANET

Definition 1: An $\langle n, r, l, m\rangle$-network is a MANET that possesses the following properties:

- The network consists of n nodes placed in an $l \times m$ rectangle area.
- The position of each node is a random variable uniformly distributed over the given area.
- Each node has a transmission radius of a uniform length r.
- A link exists between two nodes that are within the transmission range of each other[1].

A wireless node is said to cover a region if every point in this region is within the node's radio transmission range. A node placed near system boundary will cover less system area than expected, as part of its coverage region is outside the system. This is referred to as border effects. To avoid clumsy results brought by border effects, we use torus convention [8], which turns the rectangle area into a torus such that the region covered by any node is considered completely within the system. Torus convention leads to the following property.

Lemma 1: The link probability (namely, the probability of occurrence of any link) in an $\langle n, r, l, m\rangle$-network with torus convention is $p = \pi r^2/lm$ when $r \leq \min(l/2, m/2)$.

We must further restrict r's maximum value to $\min(l/3, m/3)$ when torus convention is used. The reason is that two nodes that are not neighbors but have a common neighbor can be distanced up to $2r$ from each other. When torus convention is used and the distance between them is only slightly less than $2r$, they may be incorrectly recognized as neighbors on the opposite direction if $r > \min(l/3, m/3)$, making our analysis imprecise.

The following two lemmas are essential in our derivation.

Lemma 2: [9] Given m random variables R_i, where $i = 1$ to m, $E[R_1+R_2+\cdots+R_m] = E[R_1]+E[R_2]+\cdots+E[R_m]$ regardless whether R_i's are independent to each other.

Lemma 3: The expected area jointly covered by two neighboring nodes is

$$r^2\left(\pi - \frac{3\sqrt{3}}{4}\right).$$

Proof: See Appendix. ∎

Given any node A with $m \geq 2$ neighbors, let $N(A) = \{X_1, X_2, \cdots, X_m\}$ be the set of A's neighbors. For any $X_i \in N(A)$, let $N(A)_i = \{X_j | X_j \in N(A) \wedge X_j \in N(X_i)\}$ be the set of nodes that are both neighbors of A and X_i. Note

[1] This is a simplified model as only path loss is taken into account.

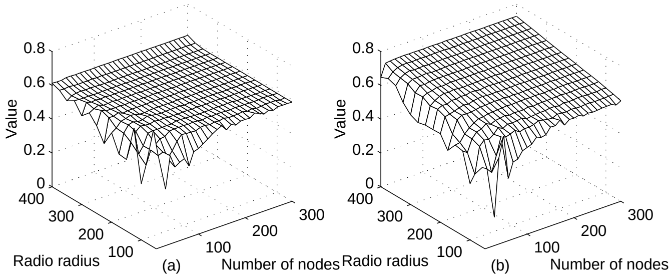

Fig. 1. Measured cluster coefficients in 1000×1000 rectangle (a) with torus convention and (b) without torus convention. Each value is averaged over 100 experiments. Nodes having less than two neighbors are not taken into account.

that $|N(A)_i|$ stands for the number of links connecting two neighbors of A such that one end of these links is X_i. The expected number of links connecting any two neighbors of A is

$$\frac{1}{2}E\left[\sum_{i=1}^{m}|N(A)_i|\right].$$

The expected value is divided by two because we count every link twice (at both ends). By Lemma 2 we have

$$\frac{1}{2}E\left[\sum_{i=1}^{m}|N(A)_i|\right]=\frac{1}{2}\sum_{i=1}^{m}E\left[|N(A)_i|\right]=\frac{1}{2}\sum_{i=1}^{m}E_{A,i},$$

where $E_{A,i}$ denotes the expected value of $|N(A)_i|$. By Lemma 3, the ratio of the region jointly covered by A and X_i to A's coverage area is expected to be

$$1-\frac{3\sqrt{3}}{4\pi}.$$

It follows that

$$E_{A,i}=(m-1)\left(1-\frac{3\sqrt{3}}{4\pi}\right)$$

for any i. Therefore, the expected number of links connecting any two neighbors of A is

$$\frac{m(m-1)}{2}\left(1-\frac{3\sqrt{3}}{4\pi}\right).$$

Dividing this value by the maximum number of links (i.e. $m(m-1)/2$) yields the expected clustering coefficient.

Theorem 1: The network clustering coefficient in an $\langle n,r,l,m\rangle$-network is expected to be a constant

$$c=1-\frac{3\sqrt{3}}{4\pi}.$$

We conducted simulations to confirm the accuracy of this theorem (See Fig. 1). The measured clustering coefficient data with torus convention have mean 0.5820 (with standard deviation 0.0313), very close to the theoretical value. The clustering coefficient without torus convention is also close to the predicted value but increases slightly with r (mean = 0.6492, standard deviation = 0.0656). Observe the little raise of the measured value with torus convention when $r>\min(l/3,m/3)$.

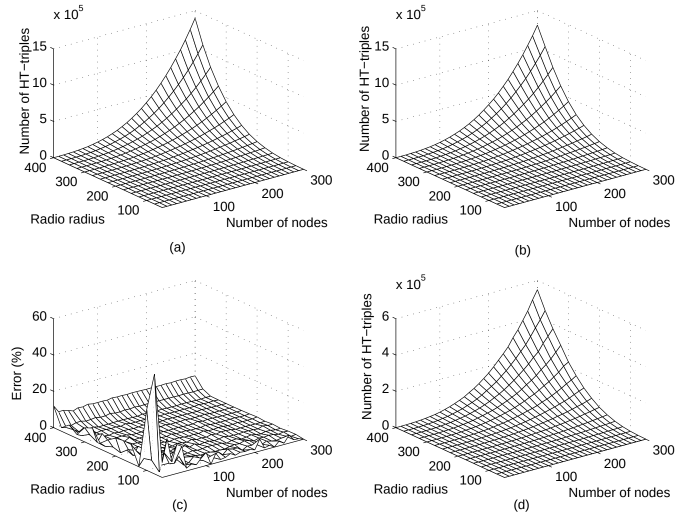

Fig. 2. Number of HT-triples in 1000×1000 rectangle. Each value is averaged over 100 experiments. (a) Theoretical result. (b) Measured result with torus convention. (c) Estimation error of (a) with respect to (b). (d) Measured result without torus convention.

III. Quantity of Hidden Terminals

Definition 2: For any three nodes X, Y, and Z, an HT-triple $\langle X,Y,Z\rangle$ is formed if both X and Z can communicate with Y but they cannot reach each other. Y is said to be the *joint node* of the HT-triple.

$\langle X,Y,Z\rangle$ forms an HT-triple if Y located within X's coverage region and Z located within Y's coverage region but not within X's. By Lemmas 1 and 3, the probability of HT-triple $\langle X,Y,Z\rangle$ is

$$\frac{\pi r^2}{lm}\times\frac{\pi r^2-r^2\left(\pi-\frac{3\sqrt{3}}{4}\right)}{lm}=(1-c)p^2. \qquad (1)$$

Theorem 2: The total number of HT-triples in an $\langle n,r,l,m\rangle$-network is expected to be

$$\eta=3\binom{n}{3}(1-c)p^2=\frac{1-c}{2}n(n-1)(n-2)p^2.$$

Proof: There are $C(n,3)$ ways to select three nodes from n nodes without order. Any selection may yield three possible HT-triples, each corresponding to a distinct joint node ($\langle X,Y,Z\rangle$ forms an HT-triple whenever $\langle Z,Y,X\rangle$ does and vise versa, so they are treated as one unique HT-triple). Although some of these HT-triples may be correlated, the expected number can still be computed (thanks to Lemma 2). ■

Note that $\eta \propto n^3p^2$. Fig. 2 compares theoretical result estimated by Theorem 2 with measured values obtained from simulations. Fig. 2(c) shows error of the theoretical estimation, where the error is defined to be

$$\frac{\text{theoretical value} - \text{measured value}}{\text{measured value}}.$$

The error is almost negligible except for the smallest n and r, where the measured value approaches zero. There is also rather high error when $r>350$ with torus convention. The measured result obtained by not using torus convention follows the same trend as the theoretical estimation, but with a different scale.

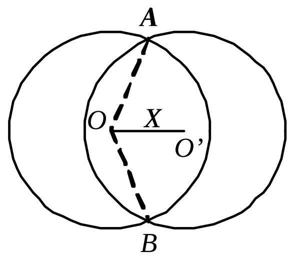

Fig. 3. Two circles intersect each other.

IV. Conclusions

We have formulated the clustering coefficient of MANETs, which turns out to be a constant with torus convention. The number of hidden terminals in a MANET is proportional to n^3p^2, where n is the number of nodes and p is the link probability. Simulation results have confirmed the accuracy of our computation.

References

[1] B. Bollobás, *Random Graphs*, 2nd ed. Cambridge University Press, 2001.

[2] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, pp. 440–442, June 1998.

[3] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74, pp. 47–97, Jan. 2002.

[4] F. A. Tobagi and L. Kleinrock, "Packet switching in radio channels: Part II - the hidden terminal problem in carrier sense multiple-access modes and the busy-tone solution," *IEEE Trans. Commun.*, vol. COM-23, no. 12, pp. 1417–1433, 1975.

[5] C. L. Fullmer and J. Garcia-Luna-Aceves, "Solutions to hidden terminal problems in wireless networks," in *Proc. SIGCOMM*, Cannes, France, 1997, pp. 39–49.

[6] P. Karn, "MACA: A new channel access method for packet radio," in *Proc. of ARRL/CRRL Amateur Radio 9th Computer Networking Conference*, Sept. 1990, pp. 134–140.

[7] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, "MACAW: A media access protocol for wireless LAN's," in *Proc. of ACM SIGCOMM'94*, London, England, 1994, pp. 212–225.

[8] P. Hall, *Introduction to the Theory of Coverage Processes*. John Wiley and Sons, 1988.

[9] E. R. Dougherty, *Probability and Statistics for the Engineering, Computing, and Physical Sciences*. Prentice-Hall, Inc., 1990, p. 238.

Appendix

Suppose that two nodes of transmission radius r located at O and O' are neighbors, with the distance between them $X \leq r$ (X is a random variable). We want to calculate the expected area of the lens-shaped region that is jointly covered by these two nodes. Let A and B be two distinct intersecting points of these two circles (refer to Fig. 3). The area of each half of the "lens" is equal to the area of sector OAB minus the area of triangle OAB. Let $\theta = \angle AOB$ be the central angle given X, where $2\pi/3 \leq \theta \leq \pi$. We have

$$X = 2r\cos(\theta/2).$$

The area of triangle OAB is

$$\left[\frac{2r\sin\left(\frac{\theta}{2}\right)\frac{X}{2}}{2}\right] = r^2\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) = \frac{r^2\sin\theta}{2}.$$

So the area of the lens is

$$2\left[\frac{\pi r^2\theta}{2\pi} - \frac{r^2\sin\theta}{2}\right] = r^2(\theta - \sin\theta).$$

Let $F(x)$ be the probability distribution function (p.d.f.) of X. Since nodes are uniformly distributed, $\Pr[X \leq x]$ is proportional to the area of the circle having radius x and being centered at O. Therefore,

$$F(x) = \Pr[X \leq x] = \frac{\pi x^2}{\pi r^2} = \frac{x^2}{r^2}.$$

Since $\theta = 2\arccos(X/2r)$, the p.d.f. of θ is

$$\begin{aligned} G(y) &= \Pr\left[\frac{2\pi}{3} \leq \theta \leq y\right] \\ &= \Pr\left[2r\cos\frac{y}{2} \leq X \leq r\right] \\ &= F(r) - F\left(2r\cos\frac{y}{2}\right) = -2\cos y - 1. \end{aligned}$$

It follows that the probability density function of θ is $g(y) = G'(y) = 2\sin y$. Therefore, the expected area of the lens-shaped region that is jointly covered by O and O' is

$$\int_{\frac{2\pi}{3}}^{\pi} r^2(\theta - \sin\theta)(2\sin\theta)d\theta = r^2\left(\pi - \frac{3\sqrt{3}}{4}\right).$$