Exact analytical solution of average path length for Apollonian networks

Zhongzhi Zhang\textsuperscript{1,2}, Lichao Chen\textsuperscript{1,2}, Shuigeng Zhou\textsuperscript{1,2}, Lujun Fang\textsuperscript{1,2}, Jihong Guan\textsuperscript{3}, and Tao Zou\textsuperscript{1,2}

\textsuperscript{1}Department of Computer Science and Engineering, Fudan University, Shanghai 200433, China
\textsuperscript{2}Shanghai Key Lab of Intelligent Information Processing, Fudan University, Shanghai 200433, China and
\textsuperscript{3}Department of Computer Science and Technology, Tongji University, 4800 Cao’an Road, Shanghai 201804, China

(Dated: February 1, 2008)

PACS numbers: 89.75.Hc, 89.75.Da, 02.10.Ox, 05.10.-a

With the help of recursion relations derived from the self-similar structure, we obtain the exact solution of average path length, $d$, for Apollonian networks. In contrast to the well-known numerical result $d \propto (\ln N_t)^{1/4}$ [Phys. Rev. Lett. \textbf{94}, 018702 (2005)], our rigorous solution shows that the average path length grows logarithmically as $d \propto \ln N_t$ in the infinite limit of network size $N_t$. The extensive numerical calculations completely agree with our closed-form solution.

PACS numbers: 89.75.Hc, 89.75.Da, 02.10.Ox, 05.10.-a

One of the most important properties of complex networks is average path length (APL), which is the mean length of the shortest paths between all pairs of vertices (nodes)\textsuperscript{1}. Most real networks have been shown to be small-world or ultra small-world networks\textsuperscript{2, 3, 4, 5}, that is, their APL $d$ behaves a logarithmic or double logarithmic scaling with the network size $N$: $d \sim \ln N$\textsuperscript{6} or $d \sim \ln \ln N$\textsuperscript{7}. It has been established that APL is relevant in many fields regarding real-life networks. In the design or interpretation of routes in architectural design, signal integrity in communication networks, the propagation of diseases or beliefs in social networks or of technology in industrial networks, APL is a natural network statistic to compute and interpret. It is strongly believed that many processes such as routing, searching, and spreading become more efficient when APL is smaller. So far, much attention has been paid to the question of APL\textsuperscript{8, 9, 10, 11, 12, 13}.

Recently, on the basis of the well-known Apollonian packing\textsuperscript{14}, Andrade et al. introduced Apollonian networks\textsuperscript{15} which were also proposed by Doye and Massen in Ref.\textsuperscript{16} simultaneously. Apollonian networks belong to a deterministic growing type of networks, which have drawn much attention from the scientific communities and have turned out to be a useful tool\textsuperscript{17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31}. Many topological properties of Apollonian networks such as degree distribution, clustering coefficient, and correlations have been determined analytically\textsuperscript{15, 16}, and the effects of the Apollonian networks on several dynamical models have been intensively studied, including Ising model and a magnetic model\textsuperscript{15, 32, 33, 34}. Despite the importance and usefulness of the quantity APL, there is no analytical calculations for the APL of Apollonian networks.

In this report, we derive an exact formula for the average path length characterizing the Apollonian networks. The analytic method is based on the recursive construction and self-similar structure of Apollonian networks.

Our rigorous result shows that APL grows logarithmically with the number of nodes. The obtained analytical solution modifies the previous numerical result in\textsuperscript{15}, where the authors claimed that the APL of Apollonian networks scales sub-logarithmically with network size. Our analytical technique could provide a paradigm for computing the APL of deterministic networks.

The Apollonian network, denoted as $A_t$ ($t \geq 0$) after $t$ generations, is constructed as follows\textsuperscript{15}: For $t = 0$, $A_0$ is a triangle. For $t \geq 1$, $A_t$ is obtained from $A_{t-1}$. For each of existing triangles created at step $t - 1$, we add a new vertex and join it to all the vertices of this triangle. Alternatively, Apollonian network can be also created in another method\textsuperscript{35, 36}. Given the generation $t$, $A_{t+1}$ may be obtained by joining at four edge nodes three copies of $A_t$, see Fig. (1a). According to the latter construction algorithm, we can easily compute the total number of vertices of $A_t$ is $N_t = (3^t + 5)/2$.

Apollonian network presents the typical characteristics of real-life networks in nature and society\textsuperscript{6, 37}. It has a power-law degree distribution with exponent $\gamma = 1 + \ln 3/\ln 2$\textsuperscript{15, 16}, which belongs to the inter-
For any individual vertex with degree \( k \), its clustering coefficient \( C(k) \) is also approximately inversely proportional to its degree \( k \) as \( C(k) = (6/k) - [2/(k-1)] \). The mean value \( C \) of clustering coefficients of all vertices is very large, which asymptotically reaches a constant value 0.8284. Moreover, the network is small world, and its diameter, defined as the longest shortest path length between all pairs of vertices, increases logarithmically with the number of vertices [38]. In fact, Apollonian network can be expanded to general cases [39]. Analogously, we can construct the stochastic version of these networks [34, 41, 41, 42]. The main topological properties of the Apollonian type networks are controlled by the dimension of packings.

After introducing the Apollonian networks, we now investigate analytically the average path length. We represent all the shortest path lengths of network \( A_t \) as a matrix in which the entry \( g_{uv} \) is the shortest path length from node \( u \) to \( v \). A measure of the typical separation between two nodes in \( A_t \) is given by the average path length \( \bar{d}_t \) defined as the mean of shortest path lengths over all couples of nodes, i.e., \( \bar{d}_t = \frac{1}{N(N-1)/2} \sum_{u,v} g_{uv} \), where \( D_t = \sum_{u,v} g_{uv} \) denotes the sum of the shortest path length between two nodes over all pairs.

By the second construction, it is obvious that Apollonian network has a self-similar structure. Thus, the total distance \( D_{t+1} \) satisfies the recursion relation

\[
D_{t+1} = 3D_t + \Theta_t - 3, \tag{1}
\]

where \( \Theta_t \) is the sum over all shortest path length whose endpoints are not in the same \( A_t \) branch. The solution of Eq. (1) is

\[
D_t = 3^{t-1}D_1 + \sum_{m=1}^{t-1} (3^{t-m-1}\theta_m) - 3^{t-1}. \tag{2}
\]

Thus, all that is left to obtain \( D_t \) is to compute \( \Theta_m \).

The paths that contribute to \( \Theta_t \) must all go through at least one of the four edge nodes (\( X, Y, Z, \) and \( O \)) at which the different \( A_t \) branches (\( A_t^{(1)}, A_t^{(2)}, A_t^{(3)} \)) are connected. The analytical expression for \( \Theta_t \), named the crossing path length, can be derived as below.

Denote \( \Theta_t^{(a,b)} \) as the sum of all shortest paths with endpoints in \( A_t^{(a)} \) and \( A_t^{(b)} \). Note that \( \Theta_t^{(a,b)} \) rules out the paths with endpoint at the edge of \( A_t^{(a)} \) and \( A_t^{(b)} \). For example, each path contributed to \( \Theta_t^{(1,2)} \) should not end at node \( O \) or \( X \). Then the total sum \( \Theta_t \) is given by \( \Theta_t = \Theta_t^{(1,2)} + \Theta_t^{(1,3)} + \Theta_t^{(2,3)} \). By symmetry, \( \Theta_t^{(1,2)} = \Theta_t^{(1,3)} = \Theta_t^{(2,3)} \), so that \( \Theta_t = 3\Theta_t^{(1,2)} \), where \( \Theta_t^{(1,2)} \) is given by the sum

\[
\Theta_t^{(1,2)} = \sum_{u \in A_t^{(1)}, v \in A_t^{(2)}, u,v \neq X \cup u,v \neq O} g_{uv}. \tag{3}
\]

To calculate the crossing path length \( \Theta_t^{(1,2)} \), we classify nodes in network \( A_{t+1} \) into seven different parts according to their shortest path lengths to each of the three vertices (i.e. \( X, Y, Z \)) of the peripheral triangle \( \triangle XYZ \). Notice that vertices \( X, Y, \) and \( Z \) themselves are not classified into any of the seven parts represented as \( P_1, P_2, P_3, P_5, P_6, \) and \( P_7 \), respectively. The classification of nodes is shown in Fig. (b). For any interior node \( v \), we denote the shortest path lengths from \( v \) to \( X, Y, Z \) as \( a, b, \) and \( c \), respectively. By construction, \( a, b, c \) can differ by at most 1 since vertices \( X, Y, \) and \( Z \) are adjacent. Then the classification function \( \text{class}(v) \) of a node \( v \) is defined to be

\[
\text{class}(v) = \begin{cases} P_1 & \text{for } a < b = c, \\ P_2 & \text{for } b < a = c, \\ P_3 & \text{for } c < a = b, \\ P_4 & \text{for } a = c < b, \\ P_5 & \text{for } a = b < c, \\ P_6 & \text{for } b = c < a, \\ P_7 & \text{for } a = b = c. \end{cases} \tag{4}
\]

It should be mentioned that the definition of node classification is recursive. For instance, Class \( P_1 \) and \( P_2 \) in \( A_t^{(1)} \) belongs to Class \( P_1 \) in \( A_{t+1} \). Class \( P_2 \) and \( P_3 \) in \( A_t^{(1)} \) belongs to Class \( P_2 \) in \( A_{t+1} \). Class \( P_3 \) in \( A_t^{(1)} \) belongs to Class \( P_7 \) in \( A_{t+1} \), and Class \( P_5 \) and \( P_6 \) in \( A_t^{(1)} \) belongs to Class \( P_8 \) in \( A_{t+1} \). Since the three vertices \( X, Y, \) and \( Z \) are symmetrical, in Apollonian network we have the following equivalent relations from the viewpoint of class cardinality: Classes \( P_1, P_2, \) and \( P_3 \) are equivalent to one another, and it is the same with Classes \( P_4, P_5, \) and \( P_6 \). We denote the number of nodes in network \( A_t \) that belong to Class \( P_k \) as \( N_{t,P_k} \), the number of classes in Class \( P_2 \) as \( N_{t,P_2} \) and so on. By symmetry, we have \( N_{t,P_1} = N_{t,P_4} = N_{t,P_7} \) and \( N_{t,P_5} = N_{t,P_6} = N_{t,P_8} \). Therefore in the following computation we will only consider \( N_{t,P_1}, N_{t,P_4}, \) and \( N_{t,P_7} \). It is easy to conclude that

\[
N_t = N_{t,P_1} + N_{t,P_4} + N_{t,P_7} + N_{t,P_8} + N_{t,P_5} + N_{t,P_6} + 3 = 3N_{t,P_1} + 3N_{t,P_4} + N_{t,P_7} + 3. \tag{5}
\]

Considering the self-similar structure of Apollonian network, we can easily know that at time \( t + 1 \), the quantities \( N_{t+1,P_1}, N_{t+1,P_4}, \) and \( N_{t+1,P_7} \) evolve according to the following recursive equations

\[
\begin{align*}
N_{t+1,P_1} &= 2N_{t,P_1} + 2N_{t,P_4}, \\
N_{t+1,P_4} &= N_{t,P_4} + N_{t,P_7}, \\
N_{t+1,P_7} &= 3N_{t,P_1} + 1,
\end{align*} \tag{6}
\]

where we have used the equivalence relations: \( N_{t,P_1} = N_{t,P_4} = N_{t,P_7} \) and \( N_{t,P_5} = N_{t,P_6} = N_{t,P_8} \). For a node \( v \) in network \( A_{t+1} \), we are also interested in the smallest value of the shortest path length from \( v \) to any of the three peripheral vertices \( X, Y, \) and \( Z \). We denote the shortest distance as \( f_v \), which can be defined to be \( f_v = \min(a,b,c) \). Let \( d_{t,v} \) denote the sum of \( f_v \) of all nodes belonging to Class \( P_2 \) in network \( A_t \). Analogously, we can also define the quantities \( d_{t,P_1}, d_{t,P_2}, \) \( d_{t,P_3}, \ldots, d_{t,P_7} \). Again by symmetry, we have \( d_{t,P_1} = d_{t,P_2} = d_{t,P_3} \).
\[ d_{t,P_1} = d_{t,P_2} = d_{t,P_3}, \] and \( d_{t,P_4}, d_{t,P_5}, d_{t,P_6} \) can be written recursively as follows:
\[
\begin{align*}
    d_{t+1,P_1} &= 2d_{t,P_1} + 2d_{t,P_4}, \\
    d_{t+1,P_4} &= d_{t,P_4} + d_{t,P_5}, \\
    d_{t+1,P_5} &= 3(d_{t,P_6} + N_{t,P_6}) + 1.
\end{align*}
\] (7)

Next we begin to determine the value of the crossing path length \( \Theta_{1,2} \). Using the definition of \( f_u \), \( \Theta_{1,2} \) can be rewritten as
\[
\Theta_{1,2} = \sum_{u \in \mathcal{A}_{t}^{(1)}, v \in \mathcal{A}_{t}^{(2)}, u,v \notin X \cup \ u,v \notin O} (f_u + f_v + \delta_{uv}),
\] (8)
where \( \delta_{uv} \) can be 0 or 1 depending on whether the two paths corresponding to \( f_u \) and \( f_v \) may meet at the same vertex (\( X \) or \( O \)). If the two paths meet at the same vertex, \( \delta_{uv} = 0, \delta_{uv} = 1 \) otherwise. According to the above classification of nodes, \( \Theta_{1,2} \) can be expressed as
\[
\Theta_{1,2}^{1,2} = \sum_{t=1}^{7} \sum_{j=1}^{7} \left( N_{t-1,P_j} d_{t-1,P_j} + N_{t-1,P_j} d_{t-1,P_j} + N_{t-1,P_j} d_{t-1,P_j} \right),
\] (9)
where \( P_1 \) and \( P_2 \) are node classes of \( A_{t}^{(1)} \) and \( A_{t}^{(2)} \), respectively. The quantity \( \delta_{P_j} \) may equal to either 0 or 1, which are related to two kinds of paths: paths from nodes in \( P_1 \) to the three peripheral vertices of \( A_{t-1}^{(1)} \), and paths from nodes in \( P_2 \) to the three peripheral vertices of \( A_{t-1}^{(2)} \). If these two sorts of paths may meet at the same peripheral vertex, \( \delta_{P_j} = 0 \); If they do not meet, \( \delta_{P_j} = 1 \). Combining previous equations and results, we get the final expression for \( \Theta_{1,2}^{1,2} \),
\[
\Theta_{1,2}^{1,2} = \frac{1}{2178} \exp(-i\pi t) \left( 2^t (-62 - 16i\sqrt{2}) + 11i2^t (10i + 7\sqrt{2}) \exp \left( \frac{i\pi t}{2} \right) - 11i2^t (-10i + 7\sqrt{2}) \exp \left( \frac{3i\pi t}{2} \right) + i2^{1+t} (31i + 8\sqrt{2}) \exp(2i\pi t) + \exp(i\pi t) \right. \\
\left. (-484 + 11 \cdot 3^{2+t} + 9^{3+t} + 22 \cdot 9^{1+t}) \right)
\] (10)

Inserting Eq. (10) into Eq. (2) and using the initial condition \( D_1 = 6 \), we have
\[
D_t = \frac{1}{330(-3 + i\sqrt{2})(-3i + \sqrt{2})} \left[ 6655i + 155(-i)^t2^{2+t} + 155(-i)^t2^{2+t} + 5i(-i)^t2^{2+t} + 5i(-i)^t2^{2+t} + 31(-1)^t2^{2+t} + 133 \cdot 3^{3+t} + 40 \cdot 9^{1+t} + 55t \cdot 3^{1+t} + 55t \cdot 3^{1+t} \right].
\] (11)

Then the exactly analytic expression for average path length can be obtained as
\[
d_{\bar{t}} = \frac{4}{1815(15 + 8 \cdot 3^t + 9^t)} \left( 6655 + 5(-i)^t2^{2+t} + 5i(-i)^t2^{2+t} + 155i(-i)^t2^{2+t} + 31(-1)^t2^{2+t} + 133 \cdot 3^{3+t} + 40 \cdot 9^{1+t} + 55t \cdot 3^{1+t} + 55t \cdot 3^{1+t} \right),
\] (12)
which can be simplified according to \( t \) as follows:
\[
d_{\bar{t}} = \begin{cases} 
\frac{46655+5 \cdot 2^{2+t}+31 \cdot 3^{2+t}+133 \cdot 3^{3+t}+40 \cdot 9^{1+t}+55 \cdot 3^{1+t}}{1815(15+8 \cdot 3^t+9^t)} & (t \mod 4 = 0), \\
\frac{46655-5 \cdot 2^{2+t}+31 \cdot 3^{2+t}+133 \cdot 3^{3+t}+40 \cdot 9^{1+t}+55 \cdot 3^{1+t}}{1815(15+8 \cdot 3^t+9^t)} & (t \mod 4 = 1), \\
\frac{46655+5 \cdot 2^{2+t}+31 \cdot 3^{2+t}+133 \cdot 3^{3+t}+40 \cdot 9^{1+t}+55 \cdot 3^{1+t}}{1815(15+8 \cdot 3^t+9^t)} & (t \mod 4 = 2), \\
\frac{46655-5 \cdot 2^{2+t}+31 \cdot 3^{2+t}+133 \cdot 3^{3+t}+40 \cdot 9^{1+t}+55 \cdot 3^{1+t}}{1815(15+8 \cdot 3^t+9^t)} & (t \mod 4 = 3)
\end{cases}
\] (13)
for \( t = 0, 1, 2, 3 \mod 4 \) are given consecutively. In the infinite network size (i.e., large \( t \)), \( d_{\bar{t}} \simeq \frac{4}{15+t} \ln N_t \). Thus, the APL scales logarithmically with increasing size of the network. We have checked our analytic result against numerical calculations for different network size up to \( t = 13 \) which corresponds to \( N_{13} = 797164 \). In all the cases we obtain a complete agreement between our theoretical formula and the results of numerical investigation, see Fig. 2.

We finally make a comment on the result presented in [15]. Using a combination of numerical experiments and heuristic arguments, Ref. [15] advances claims about the interesting quantity of average path length. The authors of [15] claimed that \( d_{\bar{t}} \propto (\ln N_t)^{3/4} \). They thus concluded that ANs belong to a new class of networks which interpolates between small \( (d_{\bar{t}} \propto \ln N_t) \) and ultrasmall \( (d_{\bar{t}} \propto \ln \ln N_t) \) networks [2]. While the combinatorial arguments and simulation results are sound,
the behavior of APL as $t$ grows is quite difficult to estimate numerically. Just as it is hard to verify with direct computation that the harmonic series diverges, the experimental results in [13] underestimate the asymptotic behavior of APL. In contrast to previous claims, our obtained precise value for APL shows that the conclusion of [13] seems questionable. The self-similar structure of Apollonian network allows us to compute precisely the quantity $APL$, which is difficult to apprehend through simulation.

In conclusion, in this report we have derived analytically exact solution for the average path length of Apollonian networks which has been attracting much research interest. We found that in the infinite network size limit Apollonian networks are small-world, the APL scales logarithmically with network size. However, some authors have been proved that conventional random scale-free networks with degree exponent $\gamma < 3$ show ultra small-world property [7, 8, 9]. In future, it is interesting to do further studies to reveal this dissimilarity between Apollonian networks and those stochastic scale-free networks, which have different scaling for APL. Finally, we believe that the analytical calculation here can guide and shed light on related studies for deterministic network models.

The authors would like to thank T. Zhou and F. Comellas for many valuable suggestions. This research was supported by the National Basic Research Program of China under grant No. 2007CB310806, the National Natural Science Foundation of China under Grant Nos. 60496327, 60573183, 90612007, 60773123, and 60704044, the Postdoctoral Science Foundation of China under Grant No. 20060400162, the Program for New Century Excellent Talents in University of China (NCET-06-0376), and the Huawei Foundation of Science and Technology (YJCB20070311N).

[1] L. da. F. Costa, F. A. Rodrigues, G. Travieso, and P. R. V. Boas, Adv. Phys. 56, 167 (2007).
[2] R. Albert and A.-L. Barabási, Rev. Mod. Phys. 74, 47 (2002).
[3] S. N. Dorogovtsev and J.F.F. Mendes, Adv. Phys. 51, 1079 (2002).
[4] M. E. J. Newman, SIAM Review 45, 167 (2003).
[5] S. Boccaletti, V. Latora, Y. Moreno, M. Chavezf, and D.-U. Hwang, Ph. Rep. 424, 175 (2006).
[6] D.J. Watts and H. Strogatz, Nature (London) 393, 440 (1998).
[7] R. Cohen and S. Havlin, Phys. Rev. Lett. 90, 058701 (2003).
[8] F. Chung and L. Lu, Proc. Natl. Acad. Sci. U.S.A. 99, 15879 (2002).
[9] S. N. Dorogovtsev, J. F. F. Mendes, and A.N. Samukhin, Nucl. Phys. 653, 307 (2003).
[10] A. Fraczak, P. Fraczak, and J. A. Holyst, Phys. Rev. E 70, 056110 (2004).
[11] W. S. Lovejoy, C. H. Loch, Soc. Netw. 25, 333 (2003).
[12] J. A. Holyst, J. Sienkiewicz, A. Fraczak, P. Fraczak, and K. Suchecki, Phys. Rev. E 72, 026108 (2005).
[13] S. N. Dorogovtsev, J. F. F. Mendes, and J. G. Oliveira, Phys. Rev. E 73, 056122 (2006).
[14] D.W. Boyd, Canadian Journal of Mathematics 25, 303 (1973).
[15] J.S. Andrade Jr., H.J. Herrmann, R.F.S. Andrade and L.R. da Silva, Phys. Rev. Lett. 94, 018702 (2005).
[16] J.P.K. Doye and C.P. Massen. Phys. Rev. E 71, 016128 (2005).
[17] A.-L. Barabási, E. Ravasz, and T. Vicsek, Physica A 299, 559 (2001).
[18] S.N. Dorogovtsev, A.V. Goltsen, and J.F.F. Mendes, Phys. Rev. E 65, 066122 (2002).
[19] S. Jung, S. Kim, and B. Kahng, Phys. Rev. E 65, 056101 (2002).
[20] F. Comellas, G. Fertin and A. Raspaud, Phys. Rev. E 69, 037104 (2004).
[21] Z. Z. Zhang, L. L. Rong, and S. G. Zhou, Physica A 377, 329 (2007).
[22] E. Ravasz, A.L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási, Science 297, 1515 (2002).
[23] E. Ravasz and A.-L. Barabási, Phys. Rev. E 67, 026112 (2003).
[24] F. Comellas, J. Ozón, and J.G. Peters, Inf. Process. Lett. 76, 83 (2000).
[25] Z.Z. Zhang, L.L Rong and C.H. Guo, Physica A 363, 567 (2006).
[26] A. N. Berker and S. Ostlund, J. Phys. C 12, 4961 (1979).
[27] M. Hinczewski, Phys. Rev. E 75, 061104 (2007).
[28] Z.Z. Zhang, S. G. Zhou, and T. Zou, Eur. Phys. J. B 56, 259 (2007).
[29] L. Barrière, F. Comellas, and C. Dalfó, J. Phys. A 39, 11739 (2006).
[30] Z.Z. Zhang, S.G. Zhou, L.J. Fang, J.H. Guan, Y.C. Zhang, EPL 79, 38007 (2007).
[31] H. D. Rozenfeld, S. Havlin, and D. ben-Avraham, New J. Phys. 9, 175 (2007).
[32] R.F.S. Andrade and H.J. Herrmann, Phys. Rev. E 71, 056131 (2005).
[33] P. Lind, J.A.C. Gallas, and H.J. Herrmann, Phys. Rev. E 70, 056207 (2004).
[34] Z. Z. Zhang, L. L. Rong, and S. G. Zhou, Phys. Rev. E, 74, 046105 (2006).
[35] E. Boltt, D. ben-Avraham, New J. Phys. 7, 26 (2005).
[36] M. Hinczewski and A. N. Berker, Phys. Rev. E 73, 066126 (2006).
[37] A.-L. Barabási and R. Albert, Science 286, 509 (1999).
[38] Z.Z. Zhang, F. Comellas, G. Fertin and L.L. Rong, J. Phys. A 39, 1811 (2006).
[39] R. Mahmoodi Baram, H.J. Herrmann, and N.Rivier, Phys. Rev. Lett. 92, 044301 (2004).
[40] T. Zhou, G. Yan, and B.H. Wang, Phys. Rev. E 71, 046141 (2005).
[41] Z.Z. Zhang, L.L. Rong and F. Comellas, Physica A 364, 610 (2006).
[42] Z. Z. Zhang and S. G. Zhou, Physica A 380, 621 (2007).
