BMC Evolutionary Biology

RESEARCH ARTICLE Open Access

MC-Net: a method for the construction of phylogenetic networks based on the Monte-Carlo method

Changiz Eslahchi[1*], Mahnaz Habibi[1], Reza Hassanzadeh[1,2], Ehsan Mottaghi[1]

Abstract

Background: A phylogenetic network is a generalization of phylogenetic trees that allows the representation of conflicting signals or alternative evolutionary histories in a single diagram. There are several methods for constructing these networks. Some of these methods are based on distances among taxa. In practice, the methods which are based on distance perform faster in comparison with other methods. The Neighbor-Net (N-Net) is a distance-based method. The N-Net produces a circular ordering from a distance matrix, then constructs a collection of weighted splits using circular ordering. The SplitsTree which is a program using these weighted splits makes a phylogenetic network. In general, finding an optimal circular ordering is an NP-hard problem. The N-Net is a heuristic algorithm to find the optimal circular ordering which is based on neighbor-joining algorithm.

Results: In this paper, we present a heuristic algorithm to find an optimal circular ordering based on the Monte-Carlo method, called MC-Net algorithm. In order to show that MC-Net performs better than N-Net, we apply both algorithms on different data sets. Then we draw phylogenetic networks corresponding to outputs of these algorithms using SplitsTree and compare the results.

Conclusions: We find that the circular ordering produced by the MC-Net is closer to optimal circular ordering than the N-Net. Furthermore, the networks corresponding to outputs of MC-Net made by SplitsTree are simpler than N-Net.

Background

Phylogenetics is concerned with the construction and analysis of phylogenetic trees or networks to understand the evolution of species, populations, and individuals. Evolutionary processes such as hybridization between species, lateral transfer of genes, recombination within a population, and convergent evolution can all lead to evolutionary histories that are distinctly non-treelike. Moreover, even when the underlying evolution is tree-like, the presence of conflicting or ambiguous signals can make a single tree representation inappropriate. In these situations, phylogenetic network methods can be particularly useful.

Phylogenetic network is a generalization of phylogenetic trees that can represent several trees simultaneously. For any network construction method, the conflicting signals should be represented in the network but it is vital that the network does not depict more conflict than is found in the data. At the same time, when the data fits well to a tree, the method should return a network that is close to a tree. Recently, in addition to biology, the phylogenetic networks methods are widely used for classifying different types of data such as those finding in linguistics, music, etc. There are many different methods to construct phylogenetic trees or networks which are based on distance matrix such as ME (minimum evolution) [1], LS (least squares) [2,3], NJ (neighbor-joining) [4], AddTree [5], N-Net (neighbor-net) [6] and Q-Net [7]. All these methods are called distance-based methods.

ME is one of the most well-known methods. It was first introduced by Kidd and Sgamarella-Zonta [1]. Given a distance matrix, the ME principle consists of selecting the tree whose length (sum of its branch lengths) is minimal among all tree topologies for taxa.

* Correspondence: ch-eslahchi@sbu.ac.ir
[1]Faculty of Mathematics, Shahid Beheshti University, G.C., Tehran, Iran
Full list of author information is available at the end of the article

Comparative studies of tree-building methods show that ME generally is an accurate criterion for selecting a true tree. Nei and Rzhetsky have shown that ME principle is statistically consistent when branch lengths are assigned by ordinary least-squares (OLS) fitting [8]. In the OLS framework, we simply minimize

$$\sum_{i,j\in X} (d_{ij} - \delta_{i,j})^2,$$

where δ_{ij} is an estimation of input d_{ij} and X is the set of taxa. In fact, the main goal is to find a tree whose induced metric is closer to d_{ij}. The LS was first introduced in [2] and [3].

Nearly 20 years have passed by since the landmark paper in Molecular Biology and Evolution introducing NJ [4]. The method has become the most widely used method for building phylogenetic trees from distances. Steel and Gascuel showed that NJ is a greedy algorithm for ME principle [9]. The N-Net is a hybrid of NJ and split decomposition [10]. It is applicable to data sets containing hundreds of taxa. The N-Net is an algorithm for constructing phylogenetic networks.

Split decomposition, implemented in SplitsTree [11], decomposes the distance matrix into simple components based on weighted splits. These splits are then represented using a special type of phylogenetic network called split network. The N-Net works in a similar way: it first produces a circular ordering from distance matrix and then constructs a collection of weighted splits. Dan Levy and Lior Patcher showed that the N-Net is a greedy algorithm for the traveling salesman problem that minimizes the balanced length of the split system at every step and it is optimal for circular distance matrices [12]. Balanced minimum evolution (BME) is designed under the ME principle [13]. The BME is a special version of the ME principle where tree length is estimated by the weighted least squares [13].

In this work, we introduce MC-Net algorithm (Monte-Carlo Network algorithm) which works in a similar way: First, it finds a circular ordering for taxa, based on Monte-Carlo with simulated annealing, it then extracts splits from the circular ordering and uses non-negative least squares for weighting splits. We compare the results of the N-Net and the MC-Net for several data sets.

Preliminaries

A *split* of a given set X of taxa is a bipartition of the set X into two non-empty subsets of X. A split is called *trivial* if one of the two subsets contains only one taxon. Let T be a non-empty tree. Let the leaves of the T are labeled by the set of taxa, $X = \{x_1,...,x_n\}$. Every edge e of T defines a split $S = A|B$, where A and B are two sets of taxa contained in the two components of T - e. For example, Figure 1 shows an eight-leaf tree. Removing the edge e from the tree produces two sets of leaves

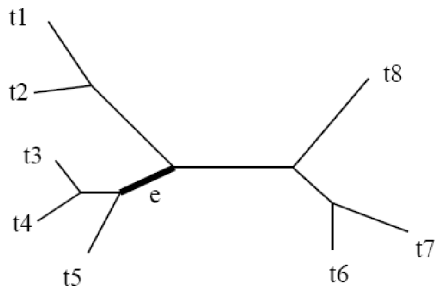

Figure 1 The split $S = A|B$ is obtained by removing the edge e of T.

$$A = \{t_3, t_4, t_5\} \text{ and } B = \{t_1, t_2, t_6, t_7, t_8\}.$$

In an edge-weighted tree, the weight of each edge is assigned to its corresponding split. The *Phyletic distance* between any two taxa x and y in an edge-weighted tree is the sum of the weights of the edges along the path from x to y. Hence, the phyletic distance between x and y equals the sum of split weights for all those splits in which x and y belong to separate components.

Two different splits $S_1 = A_1|B_1$, and $S_2 = A_2|B_2$, are *compatible*, if one of the following conditions holds:

$$A_1 \subseteq A_2, A_1 \subseteq B_2, B_1 \subseteq A_2 \text{ or } B_1 \subseteq B_2.$$

A collection of splits is called compatible, if all possible pairing of splits are compatible. A compatible collection of splits is represented by a phylogenetic tree [14,15]. Dress and Huson introduced SplitsTree to display more complex evolutionary patterns [16]. For a set of incompatible splits, SplitsTree outputs the split network using bands of parallel edges.

Circular collection of splits is a mathematical generalization of compatible collections of splits. Formally, a collection of splits of X is circular if there exists an ordering $x_1,...,x_n$ of X such that every split is of the form $\{x_i, x_{i+1},...,x_j\}|X - \{x_i,...,x_j\}$ for some i and j, $1 \le i \le j \le n$. A Compatible collection of splits are always circular [10]. On the other hand, the class of circular collection of splits contains the class of the collection of splits corresponding to a tree. Andreas Dress and Daniel Huson proved that circular collections of splits always have a planar splits graph representation [16]. A distance matrix is circular (also called Kalmanson) if it is the phyletic distances for a circular collection of splits with positive weights. Because compatible splits are circular, treelike distances are circular too [6].

As mentioned above, the ME principle consists of selecting a tree whose length is minimal. In fact, the ME

principle is equivalent to finding a circular ordering $\sigma = \{x_{\sigma(1)},...,x_{\sigma(n)}\}$ in order to find the minimum of the function η

$$\eta : \sum \to \Re$$
$$\eta(\sigma) = d(x_{\sigma(1)}, x_{\sigma(n)}) + \sum_{k=1}^{n-1} d(x_{\sigma(k)}, x_{\sigma(k+1)}) \quad (1)$$

Where Σ is the set of all circular orderings of taxa $x_1,...,x_n$. We call function η the *energy function*, and any circular ordering that minimizes η is called the *optimal circular ordering*.

Methods

There are a number of different methods for constructing various kinds of phylogenetic networks. A phylogenetic network can be constructed from a collection of weighted splits. N-Net uses circular ordering to construct a collection of weighted splits. Since finding an optimal circular ordering is an NP-hard problem, so we introduce a heuristic algorithm based on the Monte-Carlo method to find optimal circular ordering. The MC-Net seeks to find an optimal circular ordering from the distance matrix and then extracts a collection of weighted splits based on that ordering.

Algorithms

In this section, a new algorithm called the MC-Net, is presented to construct a set of weighted splits for taxa set $X = \{x_1,...,x_n\}$with a given distance matrix. The MC-Net consists of two steps. In the first step, we find a circular ordering. In the second step, the splits which are obtained from the circular ordering are weighted. The core of the first step contains two procedures, namely, INITIAL and the Monte-Carlo. The INITIAL is a greedy algorithm to obtain a circular ordering, namely, the initial circular ordering. The INITIAL works in the following way:

Suppose $x_{\sigma(1)},...,x_{\sigma(k)}$ are ordered and let $\bar{x}$ be an element of $S = X - \{x_{\sigma(1)},...,x_{\sigma(k)}\}$ such that

$$d(\bar{x}, r) = \min\{d(x, r) \mid x \in S, r \in \{x_{\sigma(1)}, x_{\sigma(k)}\}\}.$$

If $r = x_{\sigma(1)}$, we consider the new ordering $\bar{x}, x_{\sigma(1)},...,x_{\sigma(k)}$. Otherwise the ordering $x_{\sigma(1)},...,x_{\sigma(k)}, \bar{x}$ is considered. This process stops when all taxa are ordered.

The second procedure, or the Monte-Carlo procedure, relies on random iteration to find the optimal circular ordering. The Monte-Carlo algorithm starts its movement from the initial circular ordering, σ_0 For each circular ordering σ, we define the neighborhood of σ, $N(\sigma)$, by:

$$N(\sigma) = \{\tilde{\sigma} \in \sum \mid \exists k; \quad 2 \le k \le n-1; \\ \tilde{\sigma} = \{x_{\sigma(1),...,}x_{\sigma(k-1),}x_{\sigma(k+1),...,}x_{\sigma(n)}, x_{\sigma(k)}\}\},$$

where Σ is the set of all circular orderings.

We choose $\sigma_1 \in N(\sigma_0)$ randomly. if $\eta(\sigma_1) \le \eta(\sigma_0)$, then the system moves into ordering σ_1. However we allow non-greedy movements for the system in order to avoid having the system trapped in local minima. In other words, if $\eta(\sigma_1) > \eta(\sigma_0)$, then the system moves into ordering σ_1 with a small probability $e^{\frac{-\eta(\sigma_1)+\eta(\sigma_0)}{T}}$, where T is a constant temperature. For each temperature, these movements are carried out t times. To compute the minimum energy we allow this process to continue until the temperature drops to zero (see the appendix for more details). Pseudo code of the Monte-Carlo algorithm is shown in Table 1. It is noticeable that the second procedure can start from any circular ordering other than the one obtained by the INITIAL procedure.

In the final step, we use the least squares algorithm to weight the splits of obtained circular ordering. Let A be the matrix with rows indexed by pairs of taxa and columns indexed by splits. Then for each pair of texa i and j and for each split k, $A_{ij,k}$ is defined by:

$$A_{ij,k} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are on opposite of split } k; \\ 0 & \text{otherwise.} \end{cases}$$

The matrix $A = [A_{ij,k}]$ is full rank [17].

Let $d = (d_{12}, d_{13},...,d_{(n-1)n})$ be an $n(n-1)/2$ dimensional vector corresponding to rows of A where d_{ij} is obtained by distance matrix. Let b be the weight vector of splits, then the phyletic distance vector is $p = Ab$.

Table 1 Pseudo code of the Monte-Carlo algorithm with simulated annealing

Input: ***T*** **initial temperature**
σ_0 **initial ordering**
T_{low} **low temperature**
t **constant number**

$\sigma = \sigma_0$
While $T > T_{low}$
  Repeat t time
    choose random $\tilde{\sigma} \in N(\sigma)$
    If $\eta(\tilde{\sigma}) \le \eta(\sigma)$
      $\sigma = \tilde{\sigma}$
    Else
      $x = random(0, 1)$
      If $x < e^{\frac{-\eta(\tilde{\sigma})+(\sigma)}{T}}$
      $\sigma = \tilde{\sigma}$
  $T = T * 0.9$
Return σ and $\eta(\sigma)$

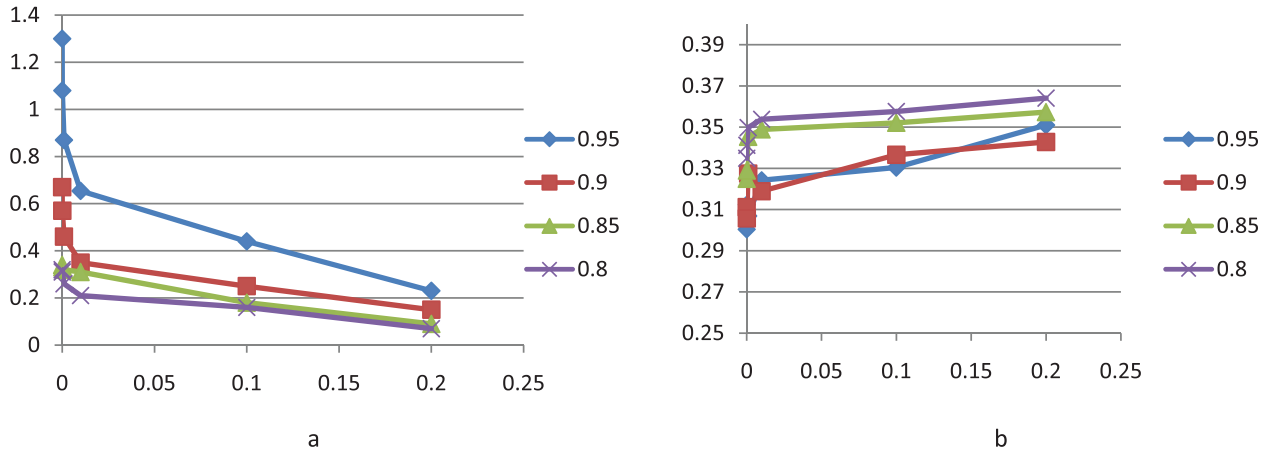

Figure 2 The value of energy function (b) and running time of algorithm (a) for each T_{low} for JSA data.

Now, the ordinary least squares(OLS) is used to estimate b by the following standard formula

$$b = (A'A)^{-1}A'd. \quad (2)$$

If we discard splits with negative weights and leave the remaining splits unchanged, the weight of the remaining splits are often grossly overestimated. Similar to the N-Net algorithm, we compute the optimal least square estimates with a non-negative constraint. In this paper, we use the FNNLS algorithm [18].

Results and Discussion

In this section, we compare the results of the MC-Net and the N-Net on some data sets. We use SplitsTree4 program [19] for drawing phylogenetic networks. Due to the limitation of space, we insert only six figures in this article.

Data sets

One of the data sets, a collection of 110 Salmonella MLST Data, was obtained from authors of the N-Net. The other data sets presented as the examples in SplitsTree4 program (version 4.10): Its(46 taxa), Jsa (46 taxa), Mammals (30 taxa), Primates (12 taxa), Rubber (23 taxa), Dolphins (36 taxa) and Myosin (143 taxa).

Table 2 Values of energy function: the values of energy function for circular orderings obtained by the N-Net, the MC-Net and the MC-Net with initial ordering of the N-Net

Data set	Its	Jsa	Mammals	Primates
N-Net	0.4096	0.2808	4.4275	2.1465
MC-Net	0.4079	0.2728	4.4172	2.1410
start N-Net	0.3979	0.2767	4.4202	2.1410

Data set	Rubber	Dolphins	Salmonella	Myosin
N-Net	0.7723	2.2	0.2546	43.8199
MC-Net	0.7596	2.1667	0.2575	43.8019
start N-Net	0.7547	2.2	0.2515	43.6935

Optimal threshold for cooling coefficient and T_{low}

There are two parameters, T_{low} and cooling coefficient, in the Monte-Carlo procedure. We first adjust T_{low} between 10^5 and 0.2 to obtain the best cooling coefficient. The value of energy function and running time of algorithm for each T_{low} for JSA data are given in Figure 2 (for the other data sets, the figures are the same as JSA). According to Figure 2, when cooling coefficient is 0.95, running time of the algorithm compared to other coefficients increases considerably. On the other hand, the value of energy function for 0.95 or 0.9 as a cooling coefficient is significantly better than the other cooling coefficients. Hence, we conclude that the best value of energy function with respect to running time of the algorithm is achieved when cooling coefficient is 0.9 and $T_{low} < 10^{-3}$.

Results

The initial test for performance of our method is done by calculating the value of energy function for circular orderings obtained by the MC-Net and the N-Net (Table 2). The first two rows of Table 2 show that in all

Table 4 The value of norm for all data sets

Data set	Its	Jsa	Mammals	Primates
N-Net	0.0444	0.0329	0.0717	0.0385
MC-Net	0.0358	0.0292	0.0648	0.0358

Data set	Rubber	Dolphins	Salmonella	Myosin
N-Net	0.0362	0.1068	0.0487	0.0291
MC-Net	0.0316	0.1019	0.0405	0.0207

Table 3 The number of splits obtained by the MC-Net and the N-Net for all data sets

Data set	Its	Jsa	Mammals	Primates
N-Net	110	83	103	34
MC-Net	105	78	99	34

Data set	Rubber	Dolphins	Salmonella	Myosin
N-Net	55	67	107	520
MC-Net	53	62	90	507

data sets except Salmonella, the value of energy function for the MC-Net is less than those obtained from the N-Net. The interesting feature of the MC-Net algorithm is in finding different circular orderings by changing initial ordering. So, the MC-Net algorithm could take the circular ordering obtained by the N-Net as initial ordering. The third row of Table 2 shows the values of energy function for circular orderings achieved by the MC-Net with the circular ordering obtained by the N-Net as an initial ordering. For four data sets, Its, Rubber, Salmonella, Myosin, the third row indicates better results than the first row. But for the other data sets, the conclusions mentioned above are the vice versa.

Another test for the performance of our method is comparing the number of splits obtained by both the algorithms. In Table 3, the number of splits of circular orderings obtained by the MC-Net and the N-Net on different data sets are shown. In all data sets the number of splits obtained by the MC-Net is less than the N-Net except Primates. In this case, these two numbers are equal.

Let d be the input distance vector and P and P' are the phyletic distance vector of weighted splits obtained by the MC-Net and the N-Net, respectively. In Table 4, the value of norm of $P - d$ and $P' - d$ for each data set

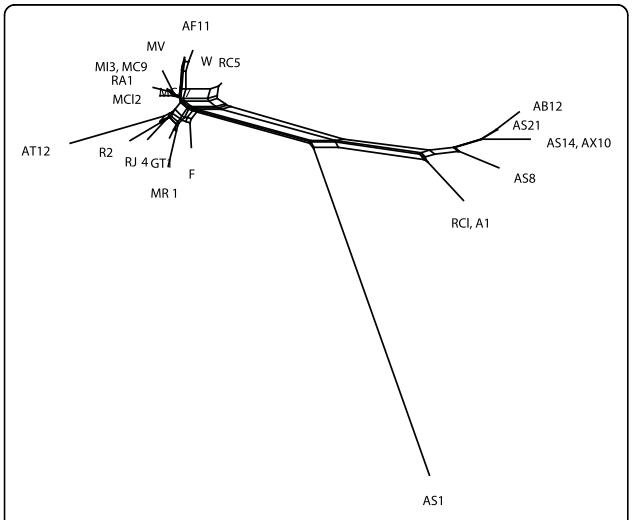

Figure 3 The N-Net network for the Rubber data set.

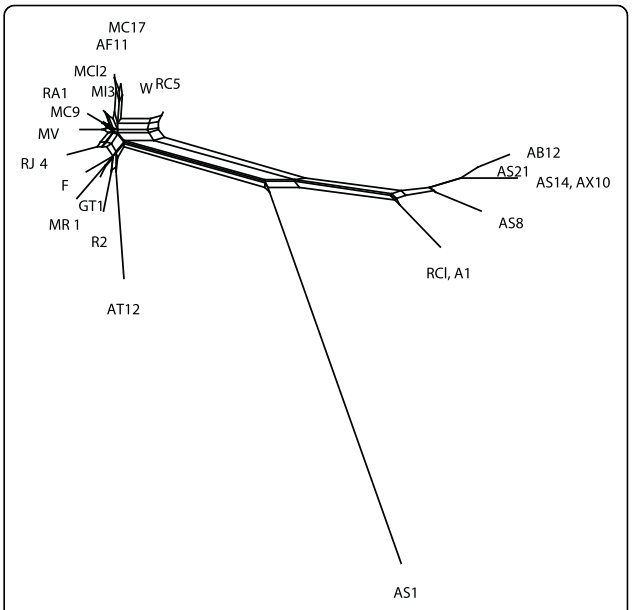

Figure 4 The MC-Net network for the Rubber data set.

are shown. The norm of $P - d$ is less than $P' - d$ in all data sets even in Primates. It means that the results of the MC-Net algorithm give better approximation for input distance vector.

To illustrate difference between two algorithms, we present some examples of networks obtained by both the MC-Net and the N-Net using SplitsTree4 (Figures 3,4,5,6, 7 and 8). It is obvious that both algorithms give the same classification of taxa and exhibit the same major splits. For example, in Figures 5 and 6, we highlight some edges such that by removing the same-colored edges, the same clustering of taxa is obtained. But according to what we see in Tables 3 and 4, split networks obtained by the MC-Net are less complicated than split networks obtained by the N-Net. It means that the networks obtained by the MC-Net have less noise than the networks obtained by the N-Net. According to Corollary 1 (see Appendix), when t approaches to 1, the MC-Net finds optimal circular ordering with the probability 1. We examined our algorithm on several treelike distance matrices and it returned corresponding trees quickly. The MC-Net has been implemented in Matlab and is available for download at http://bioinf.cs.ipm.ac.ir/softwares/mc.net.

Conclusions

In this work, we propose an algorithm, MC-Net, which is a distance based method for constructing phylogenetic networks. The MC-Net scales well and can quickly produce detailed and informative networks for large number of taxa. We compare the performance of the

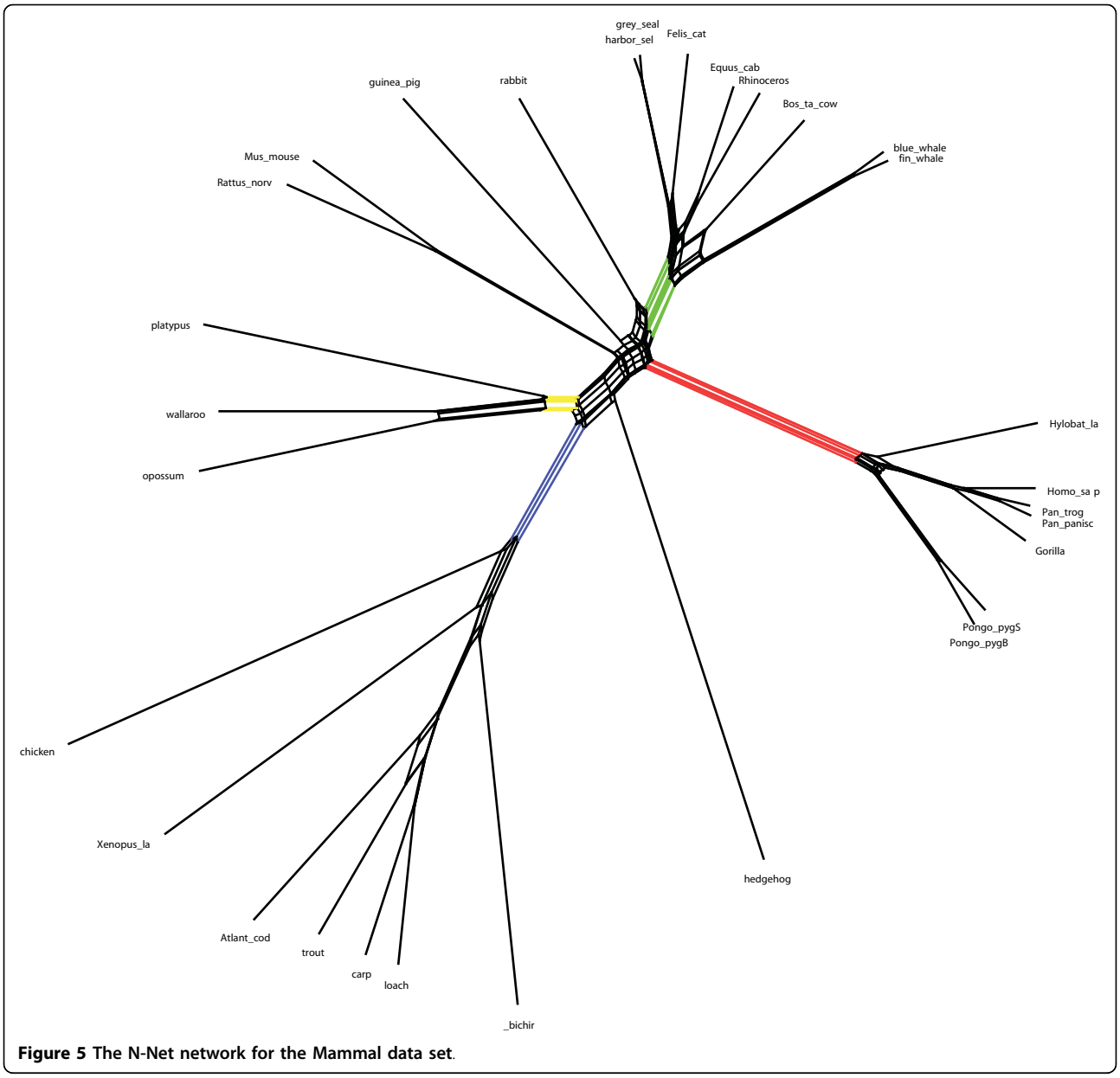

Figure 5 The N-Net network for the Mammal data set.

MC-Net with the N-Net on eight different data sets. We have shown (Tables 2, 3 and 4) that the MC-Net performs better than the N-Net for almost test cases and the networks obtained by the MC-Net are simpler than the N-Net with the same major splits. The N-Net is a part of SplitsTree program. So, the results of the MC-Net could be used in SplitsTree program too.

Appendix

Let $S = \{E_1,...,E_s\}$ be a finite set of states, and consider a physical process having these discrete states at time t. A Markov chain is a stochastic model of this system, such that the state of system at time $t + 1$ depends only on the state of system at time t.

Consider $X_0, X_1,...,$ be a collection of Markov random variables, such that X_n is the state of the system at time n. Let p_{ij} be the probability that the system enters into the state E_j from the state E_i, where $i, j \in \{1,...,s\}$ The matrix $P = (p_{ij})_{1\leq i,\ j\leq s}$ is called *transition matrix*. A probability distribution $q = (q_1,...,q_s)$ such that q_i is the probability that system starts its movement from the state E_i , is called *initial probability distribution*. A *Markov chain*

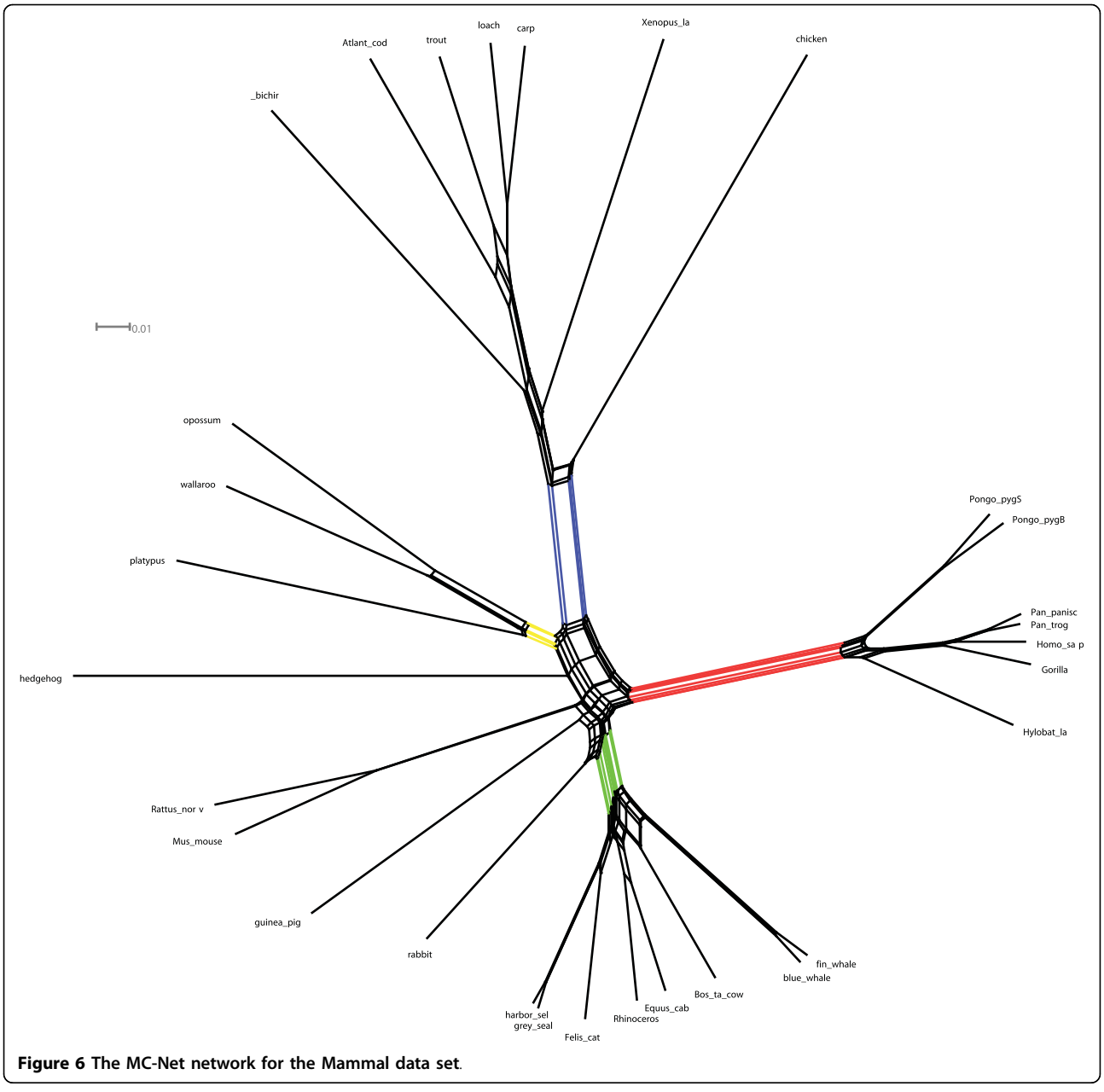

Figure 6 The MC-Net network for the Mammal data set.

is a stochastic model X_0, X_1,..., such that X_t is the state of the system at time t. For each i and j in {1,2,...,s};

$$prob\left(X_0 = E_i\right) = q_i,$$
$$prob\left(X_{t+1} = E_j \mid X_t = E_i\right) = p_{ij}.$$

The Markov chain is *irreducible*, if for all $i, j \in \{1,...,s\}$ there exists $n > 0$ such that $p_{ij}^{(n)} > 0$, where

$$\forall \alpha \quad p_{ij}^{(n)} = prob(X_{n+\alpha} = E_j \mid X_\alpha = E_i).$$

In other words, the Markov chain is irreducible, if there exist n such that the probability that the system enters into the state E_j from the state E_i after n times is positive. The irreducible Markov chain is called aperiodic, if for some $n \geq 0$ and some state E_j,

$$prob\left(X_n = E_j \mid X_0 = E_j\right) > 0$$
$$\&$$
$$prob\left(X_{n+1} = E_j \mid X_0 = E_j\right) > 0.$$

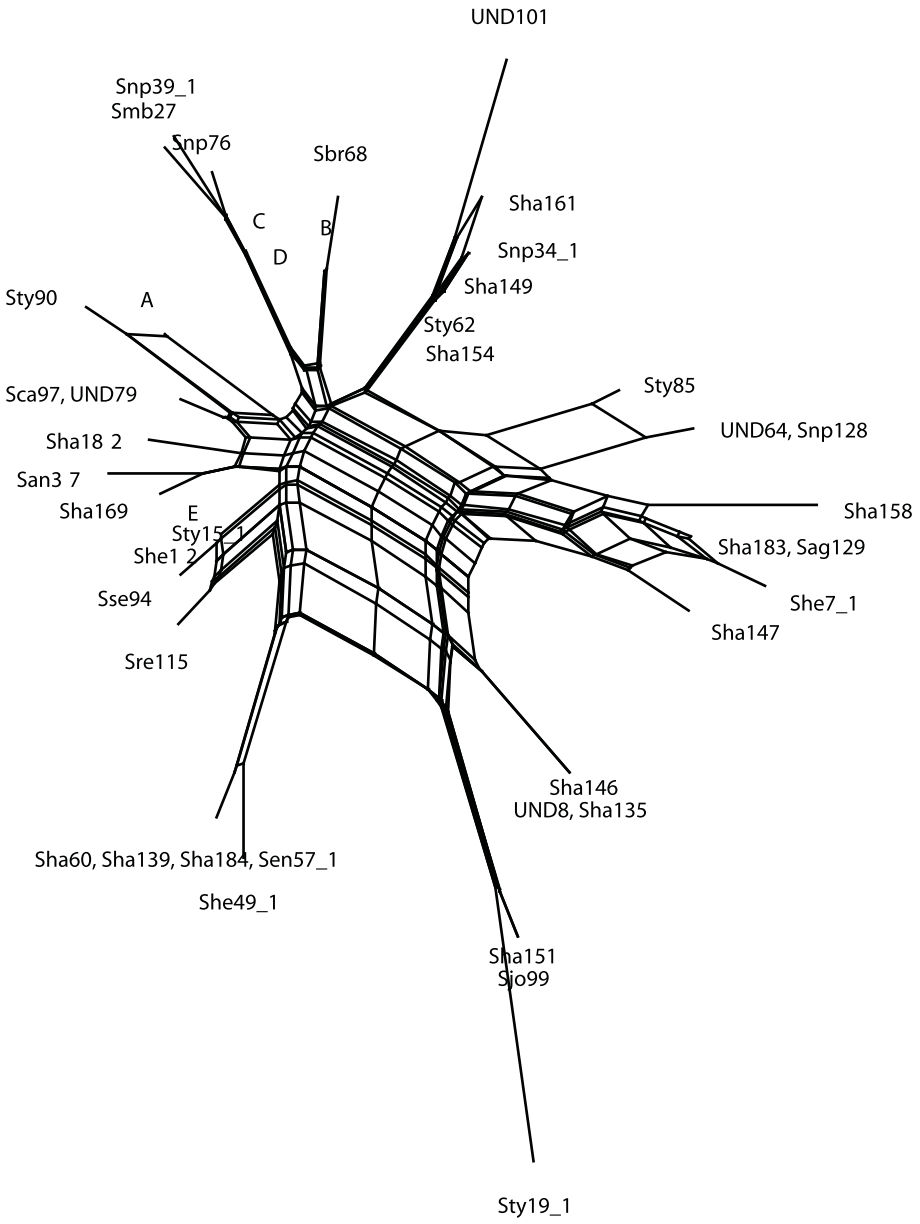

Figure 7 The N-Net network for the Salmonella data set. Group A includes the isolates Sty54, Sty54*, Sty2, She9, Sty87, Snp40*, Sty13, Snp41*, Sen5, Sha160, Sha141, Sty20*, Sha58, Sse18, Sha71, Sty31. Group B includes the isolates Sty61, Sha148, Smb-17, Sag75, Sha124. Group C includes the isolates UND3, Sha150, Sha173, Sen23*, Sha153, Sha140, San96, Sen30*, Sen24*, Sha138, Sha176, Sha130, Sha164, Sha157, Sen29*, Sca93, Sha122, Sht20, Sha186. Group D includes the isolates She3, Sha50, Sse95, Sha56, Sen24, Sen34, Sha177, Sty13*, Swo44, Sty86, Ste41, Sha77, UND80. Group E includes the isolates Ssc40, Sse28, Sty89, Sty15*, Ske69, UND110, Sha49, Sen4, Sha48, Sha165, Sty92, Snp33*, Sty52, UND109, Sha131, Sha102, Sty6, Sha175.

Theorem 1(Convergence to stationary Markov chain, [20]) If the Markov chain is irreducible and aperiodic then

$$\lim_{t\to\infty} prob\left(X_t = E_j\right) = \pi_j \qquad j = 1,\ldots,s$$

such that $\pi = (\pi_1,...,\pi_s)$ is a unique probability distribution and $\pi_j = \sum_{i=1}^{s} \pi_i p_{ij.}$.

The probability distribution is π is called *stationary probability* of the Markov chain.

It means that if P is the transition matrix and $P^{(t)}$ is the t^{th} power of P, when $t \to \infty$ the j^{th} column of transition matrix is approximately equal to π_j. In the Monte-Carlo algorithm, a special kind of Markov chain is used. Let Σ be the finite set of states and $q = \left(\frac{1}{|\Sigma|},\ldots,\frac{1}{|\Sigma|}\right)$ is

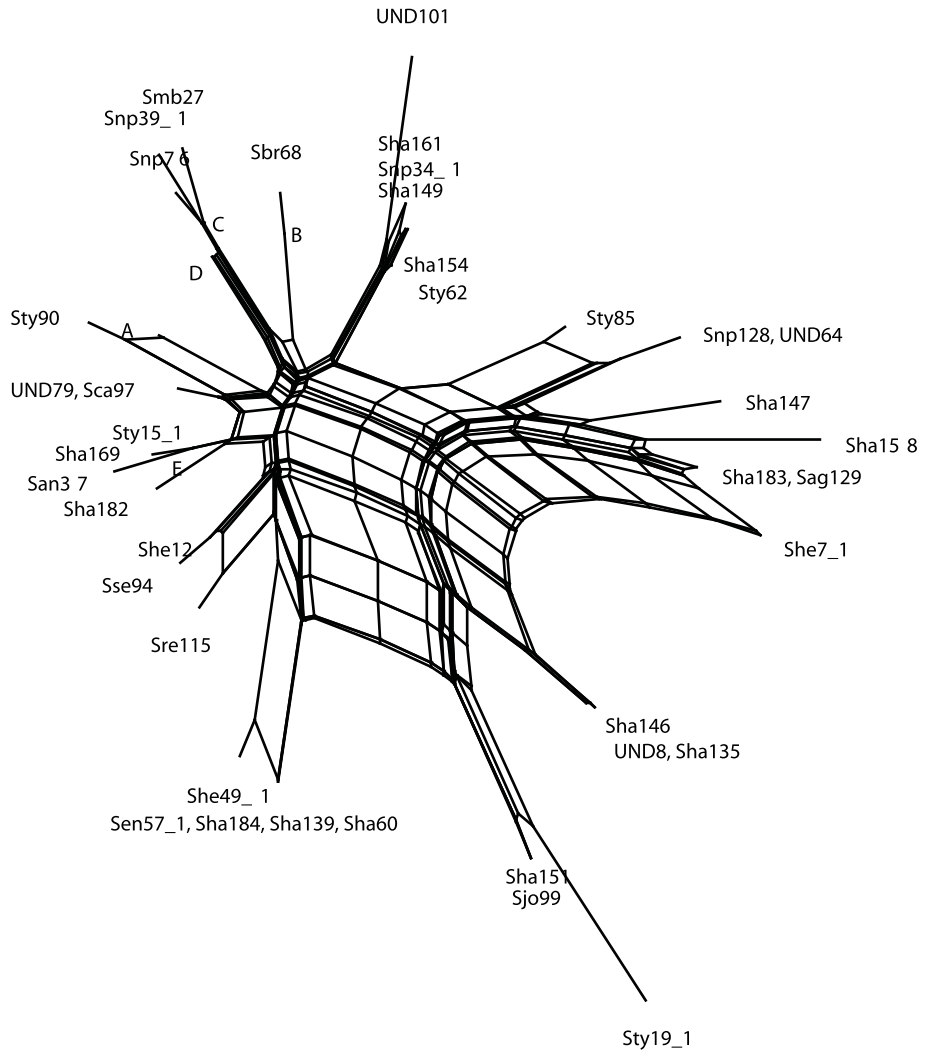

Figure 8 The MC-Net network for the Salmonella data set. Group A includes the isolates Sty54, Sty54*, Sty2, She9, Sty87, Snp40*, Sty13, Snp41*, Sen5, Sha160, Sha141, Sty20*, Sha58, Sse18, Sha71, Sty31. Group B includes the isolates Sty61, Sha148, Smb-17, Sag75, Sha124. Group C includes the isolates UND3, Sha150, Sha173, Sen23*, Sha153, Sha140, San96, Sen30*, Sen24*, Sha138, Sha176, Sha130, Sha164, Sha157, Sen29*, Sca93, Sha122, Sht20, Sha186. Group D includes the isolates She3, Sha50, Sse95, Sha56, Sen24, Sen34, Sha177, Sty13*, Swo44, Sty86, Ste41, Sha77, UND80. Group E includes the isolates Ssc40, Sse28, Sty89, Sty15*, Ske69, UND110, Sha49, Sen4, Sha48, Sha165, Sty92, Snp33*, Sty52, UND109, Sha131, Sha102, Sty6, Sha175.

the initial probability distribution. For each state i the neighborhood of i, $N(i)$, is defined as the set of all the states that are reachable from i by one movement. In this system the set of neighborhoods have to satisfy the following properties:

1. $i, \notin N(i)$.
2. $i \in N(j) \Leftrightarrow j \in N(i)$
3. if $i \neq j$, then there exit $i_1, i_2, ..., i_1 \in \Sigma$ such that

$$i \in N(i_1), i_1 \in N(i_2), \ldots, i_l \in N(j).$$

The matrix $P^T = \left(p_{ij}^T\right)_{i,j\in\Sigma}$ is defined as the transition matrix by

$$p_{ij}^T = \begin{cases} \dfrac{1}{|N(i)|} & \text{if } j \in N(i) \text{ and } \eta(j) \leq \eta(i), \\ \dfrac{e^{-(\eta(j)-\eta(i))/T}}{|N(i)|} & \text{if } j \in N(i) \text{ and } \eta(j) > \eta(i), \\ 1 - \sum_{k \in \Sigma, k \neq i} p_{ik}^T & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases}$$

where T is a positive constant number (constant temperature). The third property of the neighborhood shows that this Markov chain is irreducible. Also, if $P_{ii}^T > 0$ and P^T contains non-negative entries then $(P^T)_{ii}^{(t)} > 0$ for all $t \geq 0$. So, it is a finite, aperiodic and irreducible Markov chain. The theorem 1 shows that for each constant temperature T and $i \in \Sigma$, there exists a stationary probability distribution π_i^T such that:

$$\lim_{t\to\infty} prob(X_t = i) = \pi_i^T,$$

Where $\pi_i^T = \frac{e^{-\frac{\eta(i)}{T}}}{\sum_{j\in\Sigma} e^{-\frac{\eta(j)}{T}}}$ (see page 45 in [20]).

Proposition 1. Let $(\pi_i^T)_{i\in\Sigma}$ be a probability distribution such that:

$$\pi_i^T = \frac{e^{-\frac{\eta(i)}{T}}}{\sum_{j\in\Sigma} e^{-\frac{\eta(j)}{T}}}$$

and suppose that $m_0 = \min\{\eta(i) \mid i \in \Sigma\}$ and, $\eta_0 = \{i \in \Sigma \mid \eta(i) = m_0\}$ then for each $i \in \Sigma$, $\lim_{T\to 0^+} \pi_i^T = \pi_i^0$, where

$$\pi_i^0 = \begin{cases} \frac{1}{|\eta_0|} & \text{if } i \in \eta_0; \\ 0 & \text{otherwise.} \end{cases}$$

Proof: The proof is presented in [20] (claim 2.8 and claim 2.9).

Corollary 1. Let Σ be the finite set of states, then for each $i \in \Sigma$ we have

$$\lim_{T\to 0^+} \lim_{t\to\infty} prob(X_t = i) = \pi_i^0.$$

The corollary 1 illustrates that by cooling temperature ($T \to 0^+$), system enters into one of the states of η_0 with the probability 1 after t ($t \to \infty$) time. In this article, we define the set of all circular orderings of taxa as the finite set of states. Our definition of neighborhood in the MC-Net satisfies in three properties of neighborhood and every elements of η_0 is an optimal circular ordering. Therefore, the MC-Net yields a circular ordering with approximately minimal energy function.

Acknowledgements
We are grateful to the faculty of mathematics of Shahid Beheshti University. This work is supported in part by IPM(cs-1385-02). The authors would like to thank Prof. Hamid Pezeshk for many useful comments.

Author details
[1]Faculty of Mathematics, Shahid Beheshti University, G.C., Tehran, Iran. [2]School of Computer Science, Institute for Studies in Theoretical Physics and Mathematics (IPM), Tehran, Iran.

Authors' contributions
CE, RH and EM performed initial studies. MH designed the algorithm. RH and EM analysis the data sets. All authors participated in the writing of the manuscript. All authors read and approved the final manuscript.

Received: 9 September 2009 Accepted: 20 August 2010
Published: 20 August 2010

References

1. Kidd KK, Sgamarella-Zonta LA: **Phylogenetic analysis: concepts and methods.** *Am J Human Genetics* 1971, **23**:235-252.
2. Cavalli-Sforza LL, Edwards AWF: **Phylogenetic analysis: models and estimating procedures.** *Am J Hum Genet* 1967, **19**:233-257, (1967).
3. Fitch WM, Margoliash E: **Construction of phylogenetic trees.** *Science* 1967, **155**:279-284.
4. Saitou N, Nei N: **The neighbor joining method: a new method for reconstructing phylogenetic trees.** *Molecular Biology and Evolution* 1987, **4**:406-425.
5. Sattath S, Tversky A: **Phylogenetic similarity trees.** *Psychometrika* 1977, **42**:319-345.
6. Bryant D, Moulton V: **NeighborNet: An agglomerative method for the construction of planar phylogenetic networks.** *Molecular Biology And Evolution* 2004, **21**:255-265.
7. Grunewald S, Forslund K, Dress A, Moulton V: **QNet: An agglomerative method for the construction of phylogenetic networks from weighted quartets.** *Molecular Biology and Evolution* 2007, **24**:532-538.
8. Rzhetsky A, Nei M: **Theoretical foundation of the minimum-evolution method of phylogenetic inference.** *Mol Biol Evol* 1993, **10**:1073-1095.
9. Gascuel O, Steel M: **Neighbor-joining revealed.** *Molecular Biology and Evolution* 2006, **23**:1997-2000.
10. Bandelt HJ, Dress AWM: **Split decomposition: A new and useful approach to phylogenetic analysis of distance data.** *Mol Phyl Evol* 1992, **1**:242-252.
11. Huson DH: **SplitsTree: A program for analyzing and visualizing evolutionary data.** *Bioinformatics* 1998, **14(10)**:68-73.
12. Levy D, Patcher L: **The Neighbor-Net Algorithm.** *Advances in Applied Mathematics* .
13. Desper R, Gascuel O: **The Minimum-Evolution Distance Based Approach to Phylogenetic Inference.** In *Math Evolution and Phylogeny*. Edited by: Gascuel O. Oxford Univ. Press; 2005:.
14. Semple C, Steel M: **Cyclic permutations and evolutionary trees.** *Adv Appl Math* 2004, **32(4)**:669-680.
15. Semple C, Steel M: *Phylogenetics* Oxford, UK: Oxford University Press 2003.
16. Dress A, Huson DH: **Constructing splits graphs.** *IEEE/ACM Transactions in Computational Biology and Bioinformatics* 2004, **1**:109-115.
17. Bandelt H-J, Dress A: **A canonical decomposition theory for metrics on a finite set.** *Adv Math* 1992, **92**:47-105.
18. Bro R, Jong SD: **A Fast Non-negativity-constrained Least Squares Algorithm.** *Journal of Chemometrics* 1997, **11(5)**:393-401.
19. Huson D, Bryant D: **Application of phylogenetic networks in evolutionary studies.** *Molecular Biology and Evolution* 2005, **23**:254-267.
20. Clote P, Backofen R: *Computational molecular biology* New York, WILEY 2000.

doi:10.1186/1471-2148-10-254
Cite this article as: Eslahchi *et al.*: MC-Net: a method for the construction of phylogenetic networks based on the Monte-Carlo method. *BMC Evolutionary Biology* 2010 **10**:254.