A TSP algorithm based on link degree

Guyang Yu*, Huajun Shi

Yu Guyang: Shanghai East—China Computer Co. Ltd Postgraduate student
The 32nd Research Institute of China Electronics Technology Group Corporation
Shanghai, China
Research direction: artificial intelligence (AI)

Tutor: Shi Huajun: Shanghai East—China Computer Co. Ltd researcher
The 32nd Research Institute of China Electronics Technology Group Corporation
Shanghai, China
*annode@mail.ustc.edu.cn

Abstract: The fundamentality of the traveling salesman problem (TSP) is the choice of an edge in the next step. This paper proposes a concept of link degree, which can display the potentiality of an edge to belong to the shortest Hamiltonian cycle in a more effectively manner and, on this basis, it presents a greedy algorithm for the TSP. Meanwhile, some relevant theorems and conjectures as well as some problems triggered are discussed as well.

1. Introduction
TSP is an abbreviation of the traveling salesman problem. The problem is to find the cost of the minimum path for a single traveler starting from the starting node, passing all the given demand nodes, and finally returning to the starting node, which is to find the shortest Hamiltonian cycle of the graph. The key to the TSP is to determine which city to go next, namely, the choice of an edge. This paper proposes a new measurement method for an edge --- link degree. Having no idea of the shortest Hamiltonian cycle of the graph, the potentiality that an edge belongs to the shortest Hamiltonian cycle can be determined in a more effectively way. TSP is an NPC problem, and studies on it is of great significance. Unless particularly specified, the graphs assumed and discussed in this paper are simple graphs.

2. Link degree of an edge

2.1. Proposal of a Problem
Assume that the Hamiltonian cycle lies in a graph, how can we know that an edge definitely belongs to the shortest Hamiltonian cycle of the graph? Or how can we determine that an edge cannot belong to the shortest Hamiltonian cycle of the graph? Is there a simple way to determine how possible an edge belongs to the shortest Hamiltonian cycle of the graph?

2.2. Some simple cases
Some judgments can be quickly made on some simple graphs.
As shown in Figure 1, it can be easily determined that edge CF and edge DG definitely belong to the shortest Hamiltonian cycle of the graph.

As shown in Figure 2, assumed that it is in the Euclidean space, and in line with the properties of convex polygon, it can be determined that edges AB, BD, DE, CE and AC must belong to the shortest Hamiltonian cycle, and that edges AD, AE, BC and CD cannot belong to the shortest Hamiltonian cycle.

It is not easy to determine which edge must belong to the shortest Hamiltonian cycle of the graph in Figure 3, and is edge AB or edge BC more likely to belong to the shortest Hamiltonian cycle of the graph? But is it possible to know which edge must not belong to the shortest Hamiltonian cycle of the graph?

2.3. Definition of link degree

For edges AB and AC in Fig. 3, it is hard to clarify their strengths and weaknesses if compared directly. The solution in this paper is to make comparisons between paths.

Imagine graph G (V, E), |V|=N. Assume that when someone is observing graph G, his eyesight allows him to only observe K nodes (K < N) at a time, then graph G is broken down into subgraphs in his eyes. This process is called K division of graph G.

Definition 1. K division: graph G(V, E), |V|=N, which is broken down into different K-gons (4<=K<N).
Definition 2. Effective path $EP_k$ ($K \geq 4$): a path consisting of $K$ nodes that may belong to the shortest Hamiltonian cycle of the graph.

Corollary: if $A_0A_1...A_k-1$ is an $EP_k$, it must be the shortest path from $A_0$ to $A_k-1$ passing $A_1, A_2, ..., A_{k-2}$.

As shown in Fig. 3, CBDA is an $EP_4$. It is the shortest path from C to A passing nodes B and D. There are two possible paths from C to A through nodes B and D: CBDA and CDBA. Edges CB, BD, DA, CD, and BA participated in the competition, CB, BD, and DA won, while CD and BA failed.

Set two attributes for each edge: $S_k(X)$ represents the total number of times X wins under $K$ division, and $F_k(X)$ displays the number of times X fails under $K$ division.

Since path CBDA wins, $S_4(CB), S_4(BD), S_4(DA)$ each add 1, and so do $F_4(CD)$ and $F_4(BA)$.

This paper refers to the above process as a competition.

Definition 3. A competition: For a fixed $K$ ($K \geq 4$), if the path $A_0A_1...A_{k-1}$ has a path from $A_0$ to $A_{k-1}$ through $A_1, A_2, ..., A_{k-2}$ and there is an $EP_k$, the edge $S_k(X)$ in $EP_k$ add 1, and the other participating edges $F_k(X)$ also add 1.

The win rate of edge $X$ under $K$ division is called the link degree of edge $X$.

Definition 4. Link degree: For a fixed $K$ ($K \geq 4$), calculate $S(X, K)$ and $F(X, K)$ for each edge after all competitions, and the link degree $L_k(X)$ = $S_k(X) / (S_k(X) + F_k(S))$ of edge $X$.

For Fig. 3, assuming $K = 4$, the total score of each edge can be calculated: $S_4(AB)=1, S_4(AC)=3, S_4(AD)=5, S_4(BC)=5, S_4(BD)=3, S_4(CD)=1$. Because each edge has participated in 5 times of competition, the link degree of each edge is: $L_4(AB)=0.2, L_4(AC)=0.6, L_4(AD)=1, L_4(BC)=1, L_4(BD)=0.6, L_4(CD)=0.2$.

In some sense, the link degree of an edge represents the probability that this edge belongs to the shortest Hamiltonian cycle of the graph under $K$ division.

For Fig. 3, it can be inferred that with $K = 4$, edge BC is more likely to belong to the shortest Hamiltonian cycle than edge AB.

2.4. Conditional link degree

The link degree of edge is defined in the previous section. If, under condition A, the link degree of some edge is required to be spotted, the link degree of edge $X$ under condition A is marked as $L_k(X \mid A)$.

Assume that condition A is that edge Y belongs to the shortest Hamiltonian cycle of the graph, or that edge Z does not belong to the shortest Hamiltonian cycle of the graph. Under condition A, both $S_k(X)$ and $F_k(X)$ of edge $X$ will change, which are marked as $S_k(X \mid A)$ and $F_k(X \mid A)$, respectively. Thus define $L_k(X \mid A)$ = $S_k(X \mid A) / (S_k(X \mid A) + F_k(X \mid A))$.

Definition 5. Conditional link degree: The link degree of edge $X$ under condition $A$ is

$L_k(X|A)=S_k(X|A)/(S_k(X|A)+F_k(X|A))$

3. Some theorems and conjectures

As a basic concept, link degree has some basic theorems and conjectures, and it has to be explored and studied in many aspects.

3.1. Core conjectures

The definition of link degree naturally leads to a very important core problem.

Conjecture 1: For a fixed $K$ ($K \geq 4$), if edge $X$ has the link degree of $L_k(X) = 1$, it must belong to the shortest Hamiltonian cycle of the graph.

For $K = 4$ or $K = 5$, it can be proved simply that Conjecture 1 is true. But for a general $K$ value, neither a proof nor a counterexample for Conjecture 1 can be found. This paper tends to infer that Conjecture 1 probably holds, or at least in Euclidean space. Further conclusions are called for.

For Fig. 3, because $K = 4$, it can be clearly presumed that edges AD and BC must belong to the shortest Hamiltonian cycle of the graph.
3.2. Some simple theorems
Theorem 1: Graph G (V, E), |V| = N, for a fixed K ((2K-2) <N), if edge X belongs to the shortest Hamiltonian cycle of the graph, there must be an effective path consisting of 2K-2 nodes with edge X.

Corollary 1: If the sum of the scores of edge X is SK (X) <K ((2K-2) <N), edge X cannot belong to the shortest Hamiltonian cycle of the graph.

For Fig. 3, since N = K = 4, Corollary 1 cannot be directly used, but according to the same principle, it can be inferred that edge CD definitely does not belong to the shortest Hamiltonian cycle of the graph.

Theorem 2: Suppose LK1 (X) = 1 for edge X. If K2 > K1, then LK2 (X) = 1.

Theorem 3: Suppose that LK (X) = 1 for edge X, then LK (X | A) = 1 under condition A.

The proof of the above theorems is relatively simple, so no specific proof will be made here.

4. Link degree-based TSP algorithm
Given the link degree of each edge, the probability that the edge belongs to the shortest Hamiltonian cycle can be determined. On this basis, this paper proposes a link degree-based TSP algorithm:

I: Take a suitable K (K ≥ 4) to divide the graph (It is hoped to use a smaller K to solve a large N problem), and condition A is left blank;
II: Calculate the conditional link degree LK (X | A) of all edges X in the graph;
III: If SK (X | A) <K of edge X exists, condition A should be modified and return to II;
IV: Take out and add the edge that has the greatest link degree and does not conflict with the solution set to the solution set, and modify condition A at the same time; if there are N edges in the solution set, end; otherwise return to II.

The computational complexity of this algorithm is .

In October 2010, a British study indicated that hornets flying around flowers demonstrated the ability to easily solve the “traveling salesman problem” [1]. As for how the hornet solved this problem, there is no good explanation. This paper believes that the hornet adopts a similar strategy to this algorithm when searching for the path (This will be discussed in detail in another paper).

5. Some discussions and conclusions about this algorithm
The link degree-based TSP algorithm has some interesting and very important properties, but it can also generate many changes. Next the work will be mainly carried out around the following aspects:

5.1. Proof and Research of Conjecture 1
The proof and research of Conjecture 1 is of great significance to this algorithm. We wonder if Conjecture 1 holds or holds under what conditions.

5.2. Monotonicity and Trend of Link Degree
If edge X belongs to the shortest Hamiltonian cycle, as K increases, whether the link degree LK (X) increases monotonically and how about its tendency. Similarly, if edge X does not belong to the shortest Hamiltonian cycle, as K increases, whether the link degree LK (X) decreases monotonically and how about its tendency.

If the link degree can have this characteristic, it will be of great help to find the solution of TSP.

5.3. Relationship between K and N
For graph G (V, E), |V| = N, this paper hopes to use a smaller K to divide a large N. If K is too small, the link degree cannot precisely demonstrate the overall relationship of graph G; if K is too large, the calculation amount of the algorithm will “explode”. Whether there is a certain correlation between K and N and the correlation between the two are an important aspect for future research. Preliminary experiments displayed that K = lnN is an optional solution. Certainly, there is still a lot of work to be done in this area.
5.4. Decomposition of a Graph

The ideal state of this algorithm is to calculate the link degree of each edge and to determine that some edges must belong to the shortest Hamiltonian cycle or some must not. Under this condition, continue to calculate the link degree of other edges, and thus to find that some edges must belong to the shortest Hamiltonian cycle or some must not. This situation is called a “chain reaction”.

But this type of “chain reaction” does not occur in most cases, so sometimes the graph has to be decomposed.

If the link degree of edges AB and BC is small, it is not sufficient to determine that edge AB or edge BC must not belong to the shortest Hamiltonian cycle. Whether AB and BC can be connectively considered? If the ABC path can be determined to must not belong to the shortest Hamiltonian cycle, the problem can be divided into two. AB does not belong to the shortest Hamiltonian cycle or BC does not. In this way, the difficulty of the problem can be effectively reduced, and a “chain reaction” can be caused easier.

Another case is that if there are three choices for node A: AB, AC and AD, the graph can be divided into three subgraphs for discussion. The calculation amount is not changed though, it is easier to produce a “chain reaction” for each subgraph.

5.5. Calculation of incomplete link degree

For a large K value, it is hard to calculate the link degree of each edge. But in some cases (such as convex edges), the incomplete link degree of some edges can be calculated. In such a case, how to choose is also a research direction in the future.

5.6. Analysis of the solution

If a solution is yielded with the algorithm, how to determine whether it is the shortest Hamiltonian cycle? What is the worst result of the algorithm? There is still much work to be done in this regard.

5.6.1. Quantitative Analysis

If Conjecture 1 holds, this algorithm has a natural quantitative calculation.

Assume graph G(V,E), |V|=N, for a fixed K, if the solution is U: A0A1A2…AN-1, the possibility that the solution is the shortest Hamiltonian cycle is

\[ P(U) = L(A_0A_1, K) \times L(A_1A_2, K|T_1) \times \ldots \times L(A_{N-2}A_{N-1}, K|T_{N-2}) \times L(A_{N-1}A_0, K) \]

Generally, the P(U) value is not very small, because this algorithm has excellent convergence, and the conditional link degree of many edges that appear at the end is equal to 1.

5.6.2. Qualitative analysis

Assume graph G (V, E), |V| = N, for a fixed K, the solution of the algorithm is U: A0A1A2 ... AN-1.

If L (AiAi + 1, K) is small, it can be roughly judged that the solution U is not the shortest Hamiltonian cycle of the graph.

5.7. Improvement to the solution

Assume graph G (V, E), |V| = N, for a fixed K, if the solution of the algorithm is U: A0A1A2 ... AN-1. Is it possible to make some improvements to the solution U?

One method is to find the minimum value of L (AiAi + 1, K), delete edge AiAi + 1 in the graph, and then solve the graph again. Meanwhile, several edges in solution U can be also deleted.

Additionally, given the principle that a small probability event does not occur, a threshold value v is set. If L (X, K) <v of edge X, delete edge X from the graph, and afterwards find the shortest Hamiltonian cycle among the rest edges.

Preliminary experiments showed that the two methods perform well in solving TSP.

5.8. Directed graph

The discussion in this paper is limited to undirected graphs, but all methods and conclusions can be
easily popularized to directed graphs. Preliminary experiments showed that the link degree method seems to have better convergence for directed graphs.

5.9. Hamiltonian cycle
It is an NP problem to determine whether a Hamiltonian cycle exists in graph G. This algorithm can be easily transformed into how to find a Hamiltonian cycle of a graph.

Assume that each edge has the length of 1, and this algorithm can be transformed into one for finding Hamiltonian cycle. For example, for Figure 1, we can calculate $L_4(CF) = 1$ and $L_4(DG) = 1$.

5.10. Conclusion
TSP is a classical NP problem, which is of great significance to its research. Based on the summary of previous studies, this paper first discusses the possibility of whether a certain edge belongs to the shortest Hamilton loop of the graph, and then creatively puts forward the concepts of effective path and link degree, and then puts forward a heuristic algorithm for THE TSP problem. The research work in this area is just a beginning, and a lot of work needs to be done to improve it. This paper hopes to make some useful work for TSP through the discussion and research of this algorithm.

Acknowledgments
It took me nearly two months to finish writing this paper. In this process, it brought me infinite passion and harvest in my student life. In the writing process of the paper encountered numerous difficulties and obstacles, are in the help of students and teachers through. In the school library to find information, the teacher gave me a lot of aspects of library's support and help, especially want to thank my thesis advisor - ShiHuaJun teacher, didn't he much guidance and help, for my selfless paper modifications and improvements for me, no I this paper finally completed. Here, I would like to express my heartfelt thanks to all the teachers who have instructed and helped me! At the same time, I would also like to thank the monographs of scholars quoted in this paper. Without the inspiration and help from the research results of these scholars, I would not be able to complete the final writing of this paper. So far, I would also like to thank my friends and classmates, who gave me a lot of useful materials in the process of writing my paper, and also provided enthusiastic help in the process of typesetting and writing my paper! No man is perfect without his faults. Due to my limited academic level, the paper will inevitably have shortcomings, I sincerely hope that you can criticize and correct me.

References
[1] S. Lin & B. W. Kernighan, An Effective Heuristic Algorithm for the Traveling-Salesman Problem, Operations Research 21, 498-516 (1973)
[2] Gutin and Punnen (eds), The Traveling Salesman Problem, Future Generation Computer System 11 (1995) 175-182
[3] Golden B., A Statistical Approach to the TSP, Networks, 7, 1977, 209-225
[4] Nurmi K., Traveling Salesman problem Tools for Microcomputers, Computers and Operations Research, 18, 6, 1991, 537-547
[5] Colorini A, Dorigo M. Heuristics from nature for hard combinatorial optimization problems. International Transactions in Operational Research, 1996, 3(1): 1-21
[6] Dorigo M, Gambardella L M. An ant colony system: A cooperative learning approach to the traveling salesman problem. IEEE Trans Evolutionary Computation, 1997, 1(1): 53-66
[7] Bonabeau E, Dorigo M, Theraulaz G. Inspirations on for optimization from social insect behavior. Nature, 2000, 406(6): 39-42
[8] Zbigniew Michalewicz, and David B. Fogel (2000). How to Solve It: Modern Heuristics