情报工程 ISSN: 2095-915X
第1卷 第1期 112-119 2015年2月

TECHNOLOGY INTELLIGENCE ENGINEERING
ISSN: 2095-915X Vol.1 No.1 112-119 Feb 2015

doi:10.3772/j.issn.2095-915x.2015.01.014

K最短路径算法与应用分析*

付媛，朱礼军，韩红旗

（中国科学技术信息研究所 北京 100038）

摘要：为应对日趋庞大的网络规模和复杂的约束条件给解决k最短路径问题带来的挑战，同时给k最短路径在图书情报领域的研究提供建议，对研究现状和算法成果进行总结。首先，介绍k最短路径问题的基本概念。其次，对国内外k最短路径问题的算法思想和应用领域研究进展进行梳理，算法思想包括偏离路径算法、删除路径算法、标号算法、智能算法等，主要的应用领域有交通运输、通信网络、电力系统、图书情报领域等。最后总结k最短路径待解决的问题以及未来研究的方向。

关键词：K最短路径，偏离路径算法，标号算法

Analysis of K Shortest Paths Algorithms and Applications

FU Yuan, ZHU Lijun, HAN Hongqi

(Institute of Scientific and Technical Information of China,Beijing 100038,China)

Abstract: To cope with the challenges of solving the K shortest paths problem by growing of network scale, complicated constraint conditions, and provide suggestions for future research. Firstly, the basic concept of the K shortest paths problem is introduced. Secondly, the research on algorithms and applications at home and abroad to the K shortest paths problem are summarized. Algorithms include deviated path algorithm, route deletion algorithm, labeling algorithm, and intelligent algorithm. Application areas include transportation field, communication networks and word segmentation field. Finally, the paper concludes the K shortest paths problem to be solved and the direction of future study.

Keywords: K shortest paths problem, deviating from the path algorithm, labeling algorithm

中图分类号：G35；TP301

* 基金项目：国家自然科学基金资助项目（编号：71473237）、国家“十二五”科技支撑计划“技术创新服务平台关键技术研究及应用示范”（编号：2011BAH30B01）、中国科学技术信息研究所预研基金项目“科学合作网络的不连通问题研究”（编号：YY201418）资助。

作者简介：付媛，女，(1989-)，硕士研究生，研究方向：知识工程，E-mail：fuyuan2014@istic.ac.cn。朱礼军，男，(1973-)，博士，研究员，主要方向：语义网、知识组织系统。韩红旗，男，(1971-)，博士，副研究员，研究方向：文本挖掘、科技情报分析、科研社会网络

1 引言

最短路径是运筹学中图与网络分析的一个经典问题，用于计算在所有从起点到终点的路径中，寻求权最小的路径，权最小指的是路程最短、时间最少、花费最少等 [1]。最短路径问题一般用来解决单源最短路径问题、负权最短路径问题、每对点间的最短路径问题，分别对应的著名算法有迪杰斯特拉（Dijkstra）算法、贝尔曼 - 福特（Bellman-Ford）算法和弗洛伊德（Floyd）算法 [2~4]。该问题在管道铺设、厂区布局、设备更新等众多实际问题中应用广泛。

在应用中，最短路径方法具有很大的局限性，面对许多问题时仅运用最短路径方法已经满足不了用户的需求。在实际中常遇到因某条线路修路停用，需要选择另外一套既节约时间又可行的路线，因此需要找到几条候选最短路径以备突发情况。类似情况同样出现在出行路线查询系统中，对于用户的多样化需求，比如，除了要求时间最短外，可能还要求步行路程少于某个常数，或者换乘车数少于某个值。所以当引入其他限制条件时，仅求出最短路径是无法满足限制条件的，因此要从 k 最短路径集合中选取可行方案。求出了最短路径，还需要求解次最短路径、然后第三最短路径直到第 K 条最短路径，并排出其长度增加的顺序，即为 k 最短路径问题 [5~9] 也称为广义的最短路径问题，这样可以最大程度地满足用户对不同方案的个性化的选择需求。

当前，随着移动通信技术以及复杂网络不断取得新的进展，对于优化网络路径提出了新的挑战。作为图与网络中关键问题之一，k 最短路径问题的解决给资源分配、信道设计、优化网络结构等技术带来了新的发展机遇 [10]，同时给图书情报领域的中文分词技术以及科学合作网络的研究提供了新的思路。

2 K 最短路径主要算法思想

K 最短路径问题是最短路径问题的扩展和变形。1959 年，霍夫曼 (Hoffman) 和帕夫雷 (Pavley) 在论文中第一次提出 k 最短路径问题 [11]，随着科技的进步、人工智能的发展、交通路线和网络规模的扩大和复杂化，来自不同领域的国内外专家学者对该问题进行了广泛的研究， k 最短路径问题逐渐成为了交通、物流、运筹学、管理学、计算机科学等领域的研究热点。k 最短路径问题通常包括两类：有限制的 k 最短路问题和无限制的 K 最短路问题 [12]。前者要求最短路径集合不含有回路，而后者对所求得的最短路径集合无限制 [13]。

K 最短路径算法有三种解决思路：递推法、直接法和综合法，递推法以第 i 条最短路径 P_i 为基础，求解第 (i+1) 条最短路径 P_{i+1}, 直到求出第 k 条最短路径；直接法不需要依靠已求出的最短路径，它可以直接求出 k 条最短路径；综合法通过结合多种已有算法成果的优点来解决问题。本文对常用的以及认可度较高的 k 最短路径问题算法进行概括和归纳，分类如图 1 所示。

2.1 递推法

2.1.1 偏离路径算法思想

该算法的核心思想以最短路径 P_1 或者已经求得的第 i 条最短偏离路径 P_i 为基础，在 P_i 上除了终止节点外的其他节点中确定偏离节点，构造候选路径，从候选路径集中找到最短的一条为最短偏离路径 P_{i+1}。不同的偏离路径算法确定偏离节点和偏离路径的方法不同。

该类算法中目前使用最广泛的算法是 Yen 在 1971 年提出的以其名字命名的 Yen 算法 [14]。该算法在求 P_{i+1} 时，要将 P_i 上除了终止节点外的所有节点都视为偏离节点，并计算每个偏离节点到终止节点的最短路径，再与之前的 P_i 上起始节点到偏离节点的路径拼接，构成候选路径，进而求

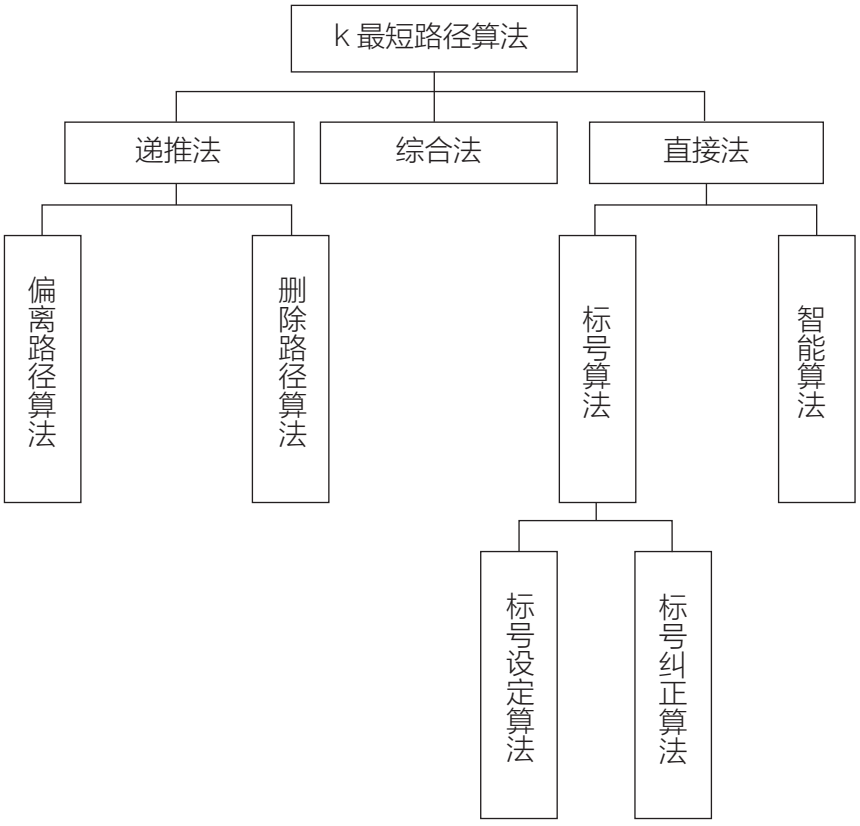

图 1 K 最短路径算法分类

得最短偏离路径。网络图中的节点数设为 n，边数设为 m，所求得路径数设为 k, 其时间复杂度为 $o(kn(m+n\log n))$。Yen 算法的优点是易于理解，可以准确地找到图中任意两节点间的 k 条最短路径，缺点是时间复杂度较高，时间代价较大，主要原因是在求 P_{i+1} 时，要将 P_i 上除了终止节点外的所有节点都视为偏离节点，从而在选择偏离节点发展候选路径时占用了大量的时间。

为提高运行速度，后人在此算法的基础上不断改进和发展。比较有效的算法之一是马丁斯（Martins）在 1999 年提出的 MPS 算法 [15]，其特点是简化了偏离路径的长度计算，在生成候选边时不像 Yen 算法那样计算每条候选路径的长度，而是要求更新每条弧上的减少长度，只选择长度减少最小的弧作为最短偏离路径，该算法在一般情况下可以提高运行速度，但是在最差的情况下与 Yen 算法的时间复杂度一样。

2.1.2 删除路径算法思想

删除路径算法 [16] 的基本思想是求解并存储从起始节点到终止节点的最短路径，然后轮流删除最短路径中的某条边，并寻找替换路径，来确定下一条最短路径，这需要通过建立附加节点和相应的边来实现。马丁斯（Martins）提出的著名的 MS 算法 [17] 就是删除路径算法，时间复杂度为 $o(m+kn\log n)$。

2.2 直接法

2.2.1 标号算法思想

该算法的核心思想是为每个节点建立一个标号链表，经过多次迭代，不断更新每个节点标号

链表以及周围序列，最终获得 k 最短路径，它是最短路径算法迪杰斯特拉（Dijkstra）算法的扩展。根据节点选择的不同方法，标号算法又分为标号设定算法和标号纠正算法 [18]。标号算法要为所有节点设置标号链表，不同点在于标号设定算法在每次迭代后，当前节点会永久标号，该节点周围关联节点会更新标号链表。标号纠正算法在每次迭代时会连当前节点及其周围关联节点的标号链表一起更新。最早的标号设置算法由德雷福斯（Dreyfus）提出，该算法用来解决最短路径问题，时间复杂度为 $o(kn^2)$。Fox 将这种思想扩展到了求解 k 最短路径问题中，时间复杂度为 $o(m+kn\log n)$。Shier 提出最早的标号纠正算法，称为双向扫除算法 [19]，该算法从前向扫除和后向扫除两个方向迭代，时间复杂度为 $o(n^3k^3)$。

2.2.2 智能算法思想

近年来，学者们从大自然生物的遗传和进化规律中得到很多启发，结合仿生学模拟动物行为建立模型，对 k 最短路径算法进行了深入的研究，获得了许多成果，提出了运用生物特性解决优化问题的新思路，比如蚁群优化算法、遗传算法等。

西安理工大学马炫 [20] 提出了求解 k 最短路径的遗传算法，该算法把问题的可行解集合映射为染色体集合的搜索空间，把节点的路径设计成染色体的编码形式，通过相应的交叉与变异方法实现了问题空间与遗传搜索空间的统一。该算法适合于大规模网络中求解 K 条最短路径。

山东大学的林洁等 [21] 提出的求解智能交通动态诱导系统中 k 最短路径问题的人工免疫搜索方法，这是一种改进的智能算法，是对生物免疫系统应用的推广和模拟并实现了并行化。算法基于生物免疫系统中抗体可以识别抗原、免疫具有记忆功能等机制，通过疫苗的提取与接种，有助于算法快速找到最优解，据此对实际网络进行建模，搜索出 K 最短路径。

2.3 综合法

综合法是近年来学者们结合已有的算法成果，引入新的思想或与其他学科结合的方法。最常见的是传统算法与启发式搜索算法相结合的方法。综合法可以结合不同算法的优点，也是未来算法一个新的研究趋势。

徐涛等 [22] 提出了约束 Yen*算法，该算法利用 Yen 算法可以准确地求出无回路的 K 最短路径问题，由于 Yen 算法运行速度较慢，所以引入了 A* 算法的启发式思想，采用 A* 算法的启发式函数，计算每个偏离节点的估价值，从中选择值最小的节点，作为发展候选路径的最终偏离节点，进而求出候选路径，缩小了求解候选路径集的搜索规模，从而提高了整体计算效率。

段宗涛等 [23] 提出了 k 最短路径集合搜索算法，该算法首先由加权邻接矩阵创建邻接表，并根据实际问题对 k 最短路径集合划分等级，创建搜索栈，使起始节点入栈，遍历查找栈顶元素的邻接节点进而访问并修改标识数组，进而求出邻接节点与栈底节点的长度值，根据路径集合选择标准来确定当前邻接节点是否入栈，直到栈顶节点为终点节点时遍历结束，搜索栈中的各元素对应的节点序列作为路径，并选择满足的最短路径集合等级。这种算法可以有效的求出各级最短路径集合，即 1 级最短路径集合，2 级最短路径集合…k 级最短路径结合。不足之处是算法效率仍然不够高，面对大规模网络无法提供实时的更新。

3 K 最短路径的应用

K 最短路径问题一直是计算机科学领域、交通运输领域的重要研究内容。它最传统的应用有出行线路的选择、交通运输方案规划以及通信线路的布置等。近年来，K 最短路径算法被用来解决一些新的实际问题，如表 1 所示。

表 1 K 最短路径应用分类

应用领域	解决的实际问题	算法
交通运输	多约束条件下快速求解国际航线网络联程路径搜索问题	约束 Yen* 算法
	求出交通网络中 k 个长度等级的最短路径集合，把大规模交通网络下的交通流量分配在不同等级的最短路径上	集合搜索算法
通信网络	在无线传感器网络的容错部署规划中，保证了无线传感器网络正常的运行且最大限度地减少了部署继电器的数量	——
	求出的 K 最大可靠路径适用于只有完好概率约束饿条件下的数据传输，求出的 K 最大容量路径适用于容量约束条件下的数据传输。	偏离路径算法
电力系统	快速对电力系统故障进行风险评估	双向搜索算法
图书情报	基于 k 最短路径的中文词语粗分方法可以找出包含正确结果在内的尽量少的 k 种粗分结果	删除路径算法

在交通运输领域的应用：徐涛等 [16] 提出约束 Yen* 算法，运用于国际航线联程路径搜索问题，在中转次数或特定中转地约束条件下，该算法采用 A* 算法的启发式思想，缩短了求解候选航线路径集合的时间，可快速查找到联程路径，解决了多约束条件下快速求解国际航线网络联程路径搜索问题，该解决方案在庞大的搜索规模下针对顾客复杂的选择心理提供了多样的选择路径。段宗涛等 [20] 提出了 k 最短路径集合算法，用来求出交通网络中 k 个长度等级的最短路径集合，并把大规模交通网络下的交通流量分配在不同等级的最短路径上，从而降低交通总代价，解决交通流合理分配问题。

在通信网络领域的应用： Lanny 等 [24] 提出将 k 最短路径运用在无线传感器网络的容错部署规划中，当网络中的最短路径故障，这时无线传感器网络重新部署规划，使次最短路启动，合理使用附加继电器的最低数量，仍然可以提供可靠的传送，保证了无线传感器网络正常的运行，最大限度地减少了部署继电器的数量。武汉大学的毛少武等 [25] 采用 MPS 算法，首先对网络中的每条弧都设定一个路径完好概率，由路径完好概率决定弧的权值，进而采用 MPS 算法求得的 k 最短路径即为路径中完好概率最大的 k 最大可靠路径；当在通信网络中传输数据量非常大时需要选取容量最大路径来传输，即为 k 最大容量路径。这时每条弧的权值为容量，并且同样采用 MPS 算法来求得 k 最大容量路径。该类算法在快速发展的通信网络，尤其是近年来兴起的物联网中有重大的应用意义。在最近的 2014 年 9 月，Justin 等 [26] 针对网络封锁问题对基于 Yen 算法的 k 最短路径的方法进行评估，评估结果表示使用 k 最短路径方

法可以生成强大的拦截方案，同时表明使用 k 最短路径可以更好地解决异构网络之间的问题。

在电力系统领域的应用：张富春等 [27] 提出利用 k 最短路径的双向搜索算法对电力系统故障进行快速风险评估，电力网中以路径的阻抗大小为权值，当某条支路断开，网络中的电流会产生剧烈变化，两点间的最短路径即为阻抗最小电流最大的路径，利用双向搜索算法确定了 k 条最短路径，进而确定了电流较大的支路，减小了风险评估的范围，提高了速度。

在图书情报领域的应用：目前主要应用于汉语词法分析方面，汉语词法分析包括：词语粗切分、切分排歧、未登录词识别、词性标注等。其中词语粗切分是一个预处理的过程，是后续语义分析的基础和前提，所以要保证粗分结果的准确性和包容性。词语粗切分的传统方法 [28] 有：全切分、基于词典的最大匹配、最短路径以及基于统计的最大概率方法等，由于中文词性的多重性，歧义现象的频发性和语言的复杂性，上述几种传统方法粗分效果较差，可能会丢失歧义词和未登录词，且效率较低。为了保证粗分结果的准确性和包容性，提高粗分效率，提出了基于 k 最短路径的中文词语粗分方法。该方法可以找出包含正确结果在内的尽量少的 k 种粗分结果，以较高的效率达到粗分的目的。

国内最早提出将 k 最短路径思想运用于汉语词法分析的是中国科学院的张华平等 [29]，提出了基于 k 最短路径方法的中文词语粗分模型，包括非统计粗分模型和基于 k 最短路径的统计粗分模型。两者的区别在于统计粗分模型比非统计粗分模型多一个步骤，要先对语料库进行统计处理，得到带有词频的概率词典，再根据概率词典构造词语切分的有向图确定每条边的权值，然后对这个构建的网络图，求出按长度值递增的 k 条最短路径，作为中文词语粗分结果集。统计粗分模型比非统计粗分模型有很大的优势，可以快速的得到数量更少、召回率更高的粗分结果集。对于小规模的语料库来说，非统计粗分模型的更为适合。目前最常用的汉语词法分析系统之一 ICTCLAS 采用的就是基于 k 最短路径的中文分词统计粗分模型切分策 [30]。

近年来有学者对基于 k 最短路径方法的中文词语粗分模型在已有的研究基础上，提出了改进方案。南京大学岑咏华 [31] 提出了基于 k 最短路径算法的中文粗分词方案，首先同样是对目标文本全切分，此方案中采用基于多重哈希词典结构进行切分，提高分词的词典匹配效率，找到所有可能的结果，然后构建有向图，基于多重哈希词典结构对有向边赋权值，采用 k 最短路径删除算法，从中找到 k 条最短路径，保留 k 个最有可能的结果，对于大规模文本，这种方法体现了较高的效率，采用的赋权机制基于最大概率模型，同时通过删除算法改进了 ICTCLAS 分词系统求解，为后续的语义分析文本处理做了良好的铺垫。李兆福 [32] 提出了求解 k 最短路径的 S-EK 最短路径算法来实现中文分词，首先通过 ICTCLAS 中文分词系统对中文字符串进行原子切分，构成特殊的有向图，对初步分词而成的 S-EK 图求解 k 最短路径。新的算法采取拆大为小、变难为易的思想，把起始点和终止点看成是两个不同的源点，图中存在这样的重要节点即起点到终点必经的节点，只需求得所有重要节点到达两边源点的 k 最短路径，提高了运行速度。

4 结束语

如今，面对瞬息万变的世界，以及用户多样化和个性化的需求，仅提供最优解已无法满足额外的限制条件，k 最短路径问题的解决为上述问题提供了优化方案。本文分析了 k 最短路径问题的经典算法以及应用的研究现状，目前已经有许

多关于 k 最短路径的算法成果，并成功的应用于实际网络中，但是其仍然具有很大的发展潜力和应用空间，还需要进一步深入研究。

首先，现有的算法研究主要趋向于对经典算法的改进和发展，并没有彻底突破传统算法，解决 KSP 问题的方法仍然比较单一，因此脱离传统算法，或者交叉结合其他思想的优势来改进传统算法将会是一个新的研究方向。

其次，对于 k 最短路径的大量研究载体都是交通网络，只有少数建立在其他网络之上，网络类型较为单一，所以算法的改进及应用研究也会具有一定的局限性。对于更具一般性的网络以及其他领域的网络，k 最短路径算法还需进一步探索。

再次，另一个被忽略的问题是，在 k 最短路径问题中，参数 k 的大小面对实际问题该如何确定，k 值的确定对算法执行效率的影响，以及它与网络大小的关系也值得进一步研究和深思，很少有学者关注过这方面的问题。

最后，上文提到的以及国内外现有的研究主要侧重于静态的以及规模固定的网络，然而，实际中许多网络的权值甚至规模都是动态变化的，因此，将各类算法合理的拓展到动态网络中或者设计适应动态网络的新算法，实现实时更新、动态变化，将是未来的趋势。同时，随着复杂网络的大规模运用和发展，动态的、复杂的网络给 k 最短路径问题的研究带来了巨大的挑战，也将带来一个新的高潮。

参考文献

[1] 王桂平 , 王衍 , 任嘉辰 . 图论算法理论、实现及应用 [M]. 北京 : 北京大学出版社 , 2011:131-208.

[2] Dijkstra E W. A note on two problems in connection with graphs [J].Numerical Mathematics, 1959, 1:269-271.

[3] 樊莉 , 孙继银 , 王勇 . 人工智能中的 A* 算法应用及编程 [J]. 微机发展 , 2003(5):33-35.

[4] 宋青， 汪小帆 . 最短路径算法加速技术研究综述 [J]. 电子科技大学学报 ,2012,41(2): 176-183.

[5] W. Xu, S. He, R. Song, et al. Finding the K shortest paths in a schedule-based transit network[J]. Computers & Operations Research, 2012, 39(8):1812-1826.

[6] 柴登峰 , 张登荣 . 前 N 条最短路径问题的算法及应用 [J]. 浙江大学学报 , 2002, 36(5):532-534.

[7] 杨旭华 , 李传告 , 陈光 . 基于 K-最短路径的交通网络传输性能分析 [J]. 浙江工业大学学报 , 2014, 42(3):250-252.

[8] Chen, Y. L. and Yang, H. H. Finding the first K shortest paths in a time window network[J]. Computers & Operations Research, 2004, 31(4):499-513.

[9] J. Hershberger, M. Maxel, S. Suri. Finding the k Shortest Simple Paths:a new algorithm and its implementation[J]. ACM Transactions on Algorithms. 2007, 3(4):45

[10] 邱钊 . K 最短路径算法及其应用研究 [D]. 成都 : 电子科技大学 , 2014:1-78.

[11] Hoffman W, Pavley R. A method for the solution of the nth best path problem[J].Journal of the Association for Computing Machinery(ACM), 1959,6(4):506-514.

[12] Eppstein D. Finding the k shortest paths[J]. SIAM Journal on computing, 1998, 28(2): 652-673.

[13] 邹永贵 , 魏来. 带多约束条件的最优路径选择算法研究 [J]. 计算机应用 , 2008, 28(5):1101-1104.

[14] Yen J Y. Finding the k shortest loopless paths in a network[J]. Management Science, 1971, 17(11): 712-716.

[15]Martins E Q V, Pascoal M M B. A new implementation of Yen's ranking loopless paths algorithm[J]. Quarterly Journal of the Belgian, French and Italian Operations Research Societies, 2003, 1(2):121-133.

[16] 高松 , 陆峰 .K 则最短路径算法效率与精度评估

[J]. 中国图象图形学报 , 2009, 14(8):1678-1682.

[17]Martins E Q V.An algorithm for ranking paths that may contain cycles[J]. European Journal of Operational Research 1984,18: 123-130.

[18] 徐涛 , 丁晓璐 , 李建伏 . K 最短路径算法综述 [J]. 计算机工程与设计 ,2013,34(11):3905-3911.

[19] Shier D R. On algorithms for finding the K shortest paths in a network[J]. Networks, 1979, 9(3): 195-214.

[20] 马炫 . 求解 k 条最优路径问题的遗传算法 [J]. 计算机工程与应用 , 2006, 42(12):100-102.

[21] 林洁 , 杨立才 , 吴晓晴 . 求解动态路径诱导 K 路最短问题的人工免疫优化方法 [J]. 山东大学学报 ,2007, 37(2): 103-107.

[22] 徐涛 , 丁晓璐 , 李建伏 . 国际航线网络联程路径搜索的 KMCSP 问题研究 [J]. 西南交通大学学报 ,2014,49(1):154-159.

[23] 段宗涛 , WANG Wei-xin, 康军 , 李莹 , 郑西彬 , 程豪 , 刘研 . 面向城市交通网络 K 最短路径集合算法 [J]. 交通运输系统工程与信息 , 2014, 14(3):195-200.

[24]Sitanayah L, Brown K N, Sreenan C J. A fault-tolerant relay placement algorithm for ensuring k vertex-disjoint shortest paths in wireless sensor networks[J]. Ad Hoc Networks, 2014, 23:145-162.

[25] 毛少武 , 张焕国 , 黄崇超 , 吴万青 . 改进的 K 最短路径算法在通信网络中的应用 [J]. 武汉大学学报 (理学版) , 2013, 59(6):534-538.

[26]Yates J, Wang X H,Chen N N. Assessing the effectiveness of k-shortest path sets in problems of network interdiction [J]. Optimization and Engineering, 2014, 15(3): 721-749.

[27] 张富春 , 黄家栋 , 周庆捷 . 电力系统连锁故障快速风险评估模型研究 [J]. 现代电力 , 2014, 31(3):1-7.

[28] 吴晓倩 , 胡学钢 . 基于 N- 最短路径的中文分词技术研究 [J]. 安徽理工大学学报 , 2014, 34(1):72-75.

[29] 张华平 , 刘群 . 基于 N- 最短路径方法的中文词语粗分模型 [J]. 中文信息学报 , 2001, 16(5):56-60.

[30] 刘群 , 张华平 , 俞鸿魁 , 程学旗 . 基于层叠隐马模型的汉语词法分析 [J]. 计算机研究与发展 , 2004, 41(8):1421-1427.

[31] 岑咏华 . 一种基于多重哈希词典和 K- 最短路径算法的中文粗分词方案研究 [J]. 情报理论与实践 , 2009 (3): 110-114.

[32] 李兆福 . 基于 K 最短路径的中文分词算法研究与实现 [D]. 哈尔滨 : 哈尔滨工程大学 , 2009, 10-50.