

上方山地区安全控制系统最优化设计

组别：第八组

指导教师：艾刚

编者：栗云昊 马众孚 李佳欣 李杰

日期：2019.07.10

目录

一、引言	1
关键字：拓扑分析、逻辑回归、方案评价.....	2
二、基于图论的拓扑结构模型	4
2.1 问题分析和模型假设.....	4
2.2 全局最优巡逻区域最大覆盖调度方法.....	5
2.3 基于遗传算法的全局巡逻部署策略.....	6
三、突发情况发生概率模型	9
3.1 基于逻辑回归的火灾概率模型.....	9
3.2 游客突发状况模型.....	13
四、突模型应用于可行性方案的提出	15
4.1 GPS 路线图.....	15
4.2 安全控制网的搭建原则.....	17
4.3 安全控制网搭建方案.....	19
4.4 方案的评价.....	21

五、模型稳定性分析22

参考文献:24

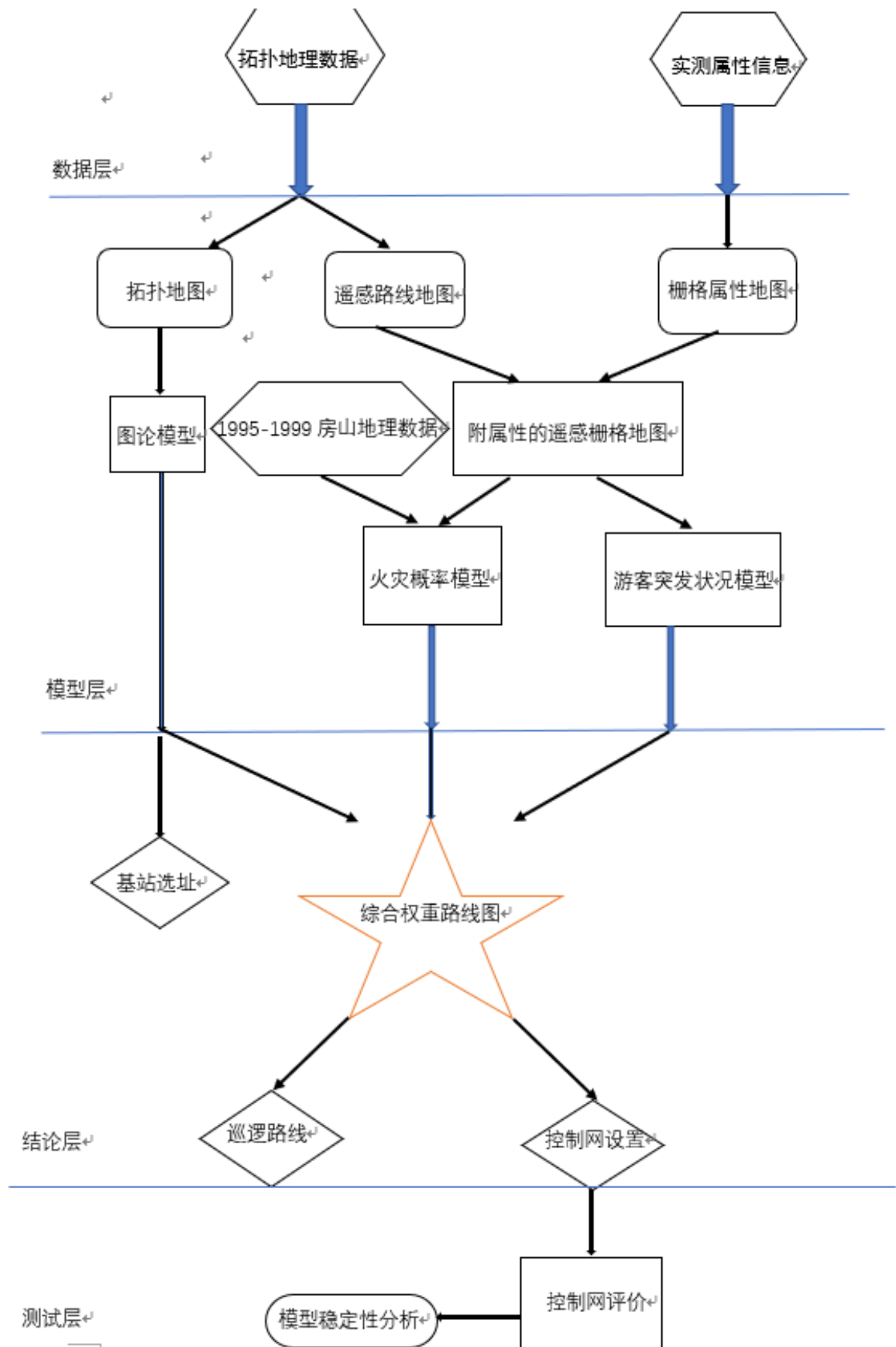
一、引言

山区安全是影响山区可持续经营的一个重要因素,对于区域山区安全防控而言,必须要有数量一定、质量可靠、结构科学的山区安全防控建设。山区火灾一旦发生,就会带来不同程度的人力、物力、财力损失,造成严重的后果,如 2019 年 4 月份雷击一棵 80 年树龄云南松造成四川凉山大火造成冷杉、云南松、杂灌等植被的烧毁及大量建筑的坍塌与山上值班人员的人身伤亡,甚至有 30 名消防人员在扑火工作中遇难。同时,山区游客遇难事件频发,如近年来在桂林市叠彩山景区内发生一起落石事故造成 7 人遇难,19 人受伤,除落石外,山区旅游还有其他危险,像蛇、猕猴、摔伤、高温、车祸、缆车意外事故等。由此,设立安保点、消防点、火灾预警装置及安保救援人员的空间合理配置十分必要。

但在此之前,可行的管理模式是一切的基础,山区开发商无法直接提供昂贵的安全器材和专业安全人员,以类型相同的泰山风景区为例,泰山的消防管理模式是以 119 消防部队为主,采用公私合营的模式。景区内安置串联式水泵、水库塘坝等水灭火设施,常驻巡逻人员防止森林火灾发生及游客遇难时及时救援。

因此参照泰山的模式,上方山也应以防火安全为主,基于山区具体干度、湿度、坡度等地形地貌因素,运用模糊层次分析、Logistic 回归、遗传算法、k-means 聚类分析等方法建立消防安保网络的合理布局,防止灾害的发生。

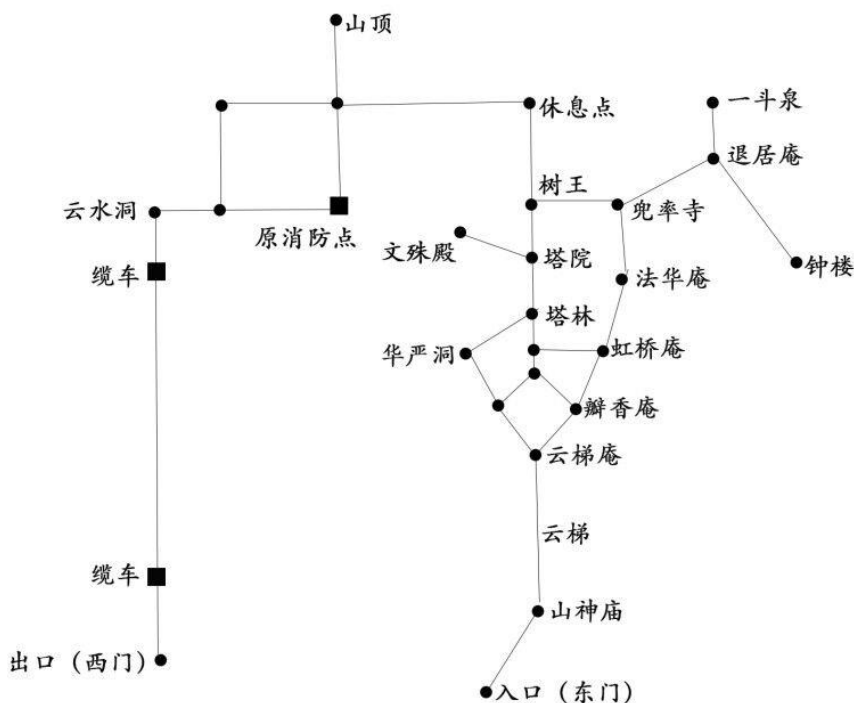
关键字: 拓扑分析、逻辑回归、方案评价



二、基于图论的拓扑结构模型

首先, 我们将上方山地图简化为拓扑图, 在以同性点为基础的简化图上使用图论知识, 先确立综合安全控制网的基础, 防火控制网的基础结构, 包括基站 的设置, 循环路线及最少循环人数的安排。

我组通过实地的调查与记录, 通过走访所有的常用路线, 简化出了全山基础 拓扑结构。得图如下:



2.1 问题分析和模型假设

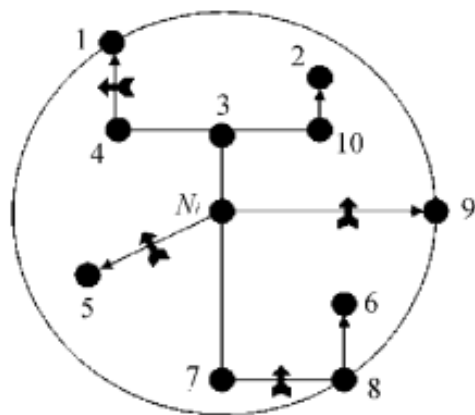
本文重点考虑某城市的巡逻队部署方案和实施策略, 共有交叉路口 26 个, 有效路线 30 条和 1 个重要点。若节点 $node1$ 和 $node2$ 是连通的, 则二者间的关联 $TEdge (node1, node2)$ 不为空, 反之为空。给出各个交叉路口和重

要点的坐标和具有连通关系的交叉路口的统计数据,并假设巡逻队的平均巡逻速度为 2m/s , 接警后的平均行驶速度达 4m/s 。图 1 给出了路口的分布, 其中符号 Δ 为重要点, 节点 \bullet 为交叉路口, 直线表示交叉路口间的连通关系。通过查询相关资料我们得知山火着火后的安全时间为 5 分钟, 因此, 方案应要求巡逻队能在接警后 5min 内赶到事故现场的比例不低于 90% , 赶到重要点的时间必须在 3min 之内, 同时考虑火灾现场的发生位置、巡逻人员的初始位置和数量的选取, 该问题是一个多目标的优化问题。在方案设计中, 将发生在道路上的火灾事故位置近似为距离该事故中心最近的节点, 最终计算出满足要求的各巡逻队的巡逻区域范围和部署路线。

2.2 全局最优巡逻区域最大覆盖调度方法

2.2.1 巡逻初始位置和数量的选取

巡逻队初始位置的选取可以看作是最小顶点覆盖问题, 要确定巡逻队的初始数量 N , 本文将问题转化为要计算出 N 辆巡逻队的巡逻区域覆盖率。将发生在道路上的事故位置近似为距离该事故中心最近的路口。



以图 1 为例, 中心的节点 \bullet 表示第 N_i 巡逻队的起始位置, 假设周围的 10 个交叉路口均可被该巡逻队覆盖到, 即巡逻队能在接警后 3min 内赶到任何现场。此时若有事故发生在道路上, 而不是发生在节点处, 如箭头所示, 则该巡逻队 N_i 同样可

在规定时间内赶到事故现场, 并成功覆盖该区域。方案中, 首先采用 *k - mean* 聚类算法对交叉路口数据进行聚类分析, 并将巡逻队的位置初始化为距离这 N_i 个中心点最近的 N_i 个交叉路口, 通过计算覆盖率确定巡逻队的初始数量。

定义 I: 若给出 N 辆警车覆盖到的路口集合

$$\bigcap_{i=1}^N T_area(i)$$

(1) 则巡逻区域覆盖率可表示为

$$P_{covered} = \left(\prod_{i=1}^N \frac{T_area(i)}{26} \right) * 100\%$$

(2) 将 $P_{covered}$ 与 0.9 比较, 若 $P_{covered}$ 小于 0.9, 则更新并循环计算, 直至满意。对于重要点的处理办法是计算 N 辆警车是否能在 2min 内赶到重要点的布尔值记为 i , 则警车数量为 $N + 1 - i$, $0 \leq i \leq 1$ 。

2.3 基于遗传算法的全局巡逻部署策略

2.3.1 警车初始化配置方案

在初始方案中, 该 N 辆警车可有效处理巡逻区域内的事故, 并能够在接警后以 $4m/s$ 的速度赶到巡逻区域内任何一个路口。在实验中, 选取 N 为 1~5 进行测试, 每组测试 50 次, 将 $P_{covered}$ 按列存储在矩阵 $M_accuracy$ 中, 将选取不同警车数量情况下的巡逻区域覆盖率的均值 $M_accuracy$ 作为实验比较的主要依据, 并对得到的数据进行统计分析, 通过计算 $M_accuracy$ 与标准 0.9 的相对误差发现, 当 $N = 2$ 时, 相对误差就能达到 **3.57%**, 即有超过 96.43% 的可能性巡逻队能够及时发现并扑灭火灾, 由此初步确定警车数量 N 取 2。此

时我们应该注意对重要点的覆盖, 使用公式 $N = 17 + 3 - i$ 表示之后遗传算法中 N 的取值, 其中 i 为重要点中是否被两队巡逻队成功覆盖。

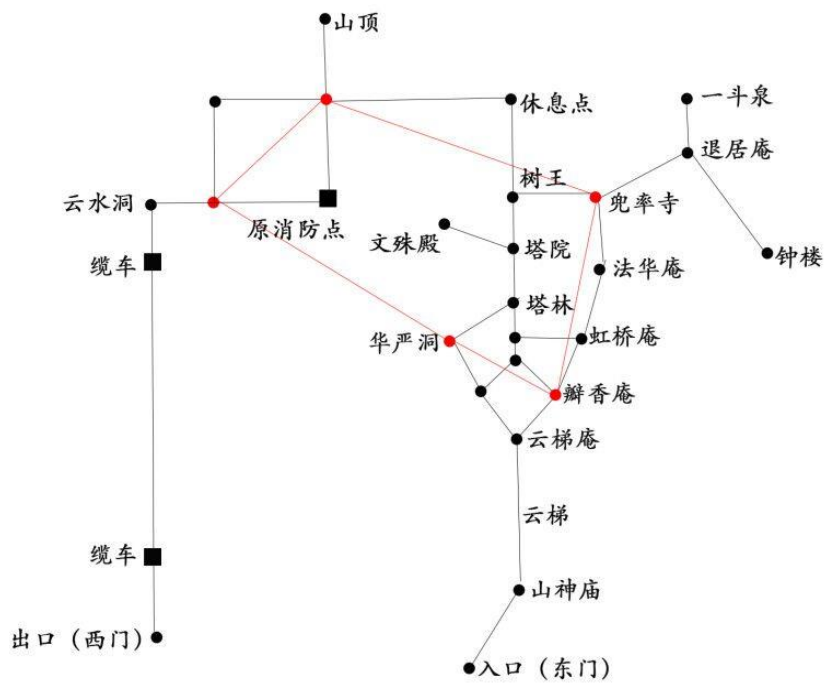
2.3.2 全局巡逻部署策略设计思路

遗传算法 [3, 5] 是一种基于自然选择与遗传机理的随机搜索算法, 为了优化警车巡逻的行车路线, 利用遗传算法优化选择全局最优路线, 且具有较强的鲁棒性。由于本次模拟的拓扑路线十分简单, 所以巡逻中, 可能会出现多辆警车重复巡逻某个交叉路口的情形, 因此可在遗传算法的优化函数中加入对重复覆盖的惩罚机制, 以保证巡逻路线不会太过复杂或有太多的重复区域

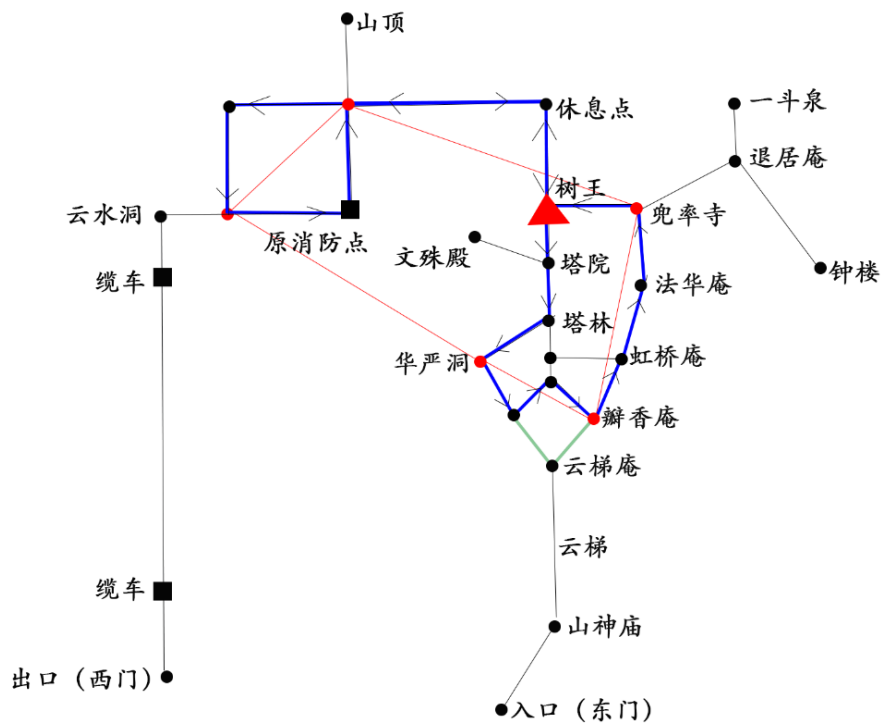
设: 两个巡逻区域交叉处的路口即为重叠遍历 $T_{area}(i)$ 和 $T_{area}(j)$ 分别为警车 N_i 和 N_j , $i \neq j$ 的巡逻区域, 交叉路口节点 5、6 和 75 可同时被警车 N_i 和 N_j 覆盖到, 被视为被重叠遍历。初始化 $POS = zeros(1, 26)$, 用于记录某路口是否已被巡逻过, 当某警车巡逻到某路口 $node$ 时, 标定 $POS(node) = 1$, 表明该路口可被覆盖, 此时若其他警车出现重复巡逻该路口时, 状态 $POS(node)$ 仍为 1, 保持不变。这样以两点最短路径与 POS 值作为优化对象, 即可防止子代的选点过于分散或过于重复。

2.3.3 实验结果

实验中, 遗传算法采用单点交叉, 交叉概率取 0.8, 变异概率取 0.2, 子群个数取 60。采用遗传算法得到代表进化 500 代后选定的最小覆盖集节点的子代, 使用 matlab 统计算法统计全部子代位置, 找出其中分布律高于普遍遗传率, 即具有遗传优势的点。画在之前的拓扑图上可得到



之后，我们就可以通过肉眼连接两点的最短路线得到真实的巡逻路线。



这样，我们从整个巡逻路线中选取中间节点，就可以找到消防基站的初步选址。但是要注意，在华严洞和瓣香庵中间存在两条最短路径，这是在单纯图论模型中无法解决的问题，我们将在此保留，并在之后的模型中进行回答。

三、突发情况发生概率模型

3.1 基于逻辑回归的火灾概率模型

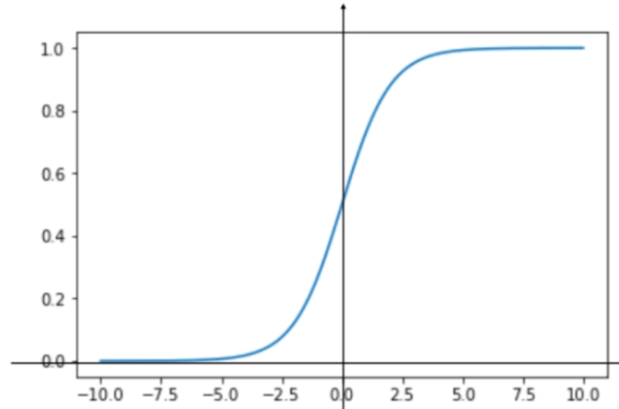
3.1.1 问题分析和模型假设

森林中的火灾的发生一定与自然因素与人为因素两种因素有关，而且所有的因素应该决定火灾发生的概率大或者小，但无论如何，火灾的发生只存在 1 和 0，即发生和未发生两种状态，这应符合逻辑回归模型。Logistic 回归 (Logistic Regression) 又称 Logistic 回归分析，是一种广义的线性回归分析模型，常用于数据挖掘，疾病自动诊断，经济预测等领域。近年来，国内外部分学者引入此方法构建森林火险模型（。许多社会现象本质上是定性的，而不是定量的，比如某个事件是否发生、个人做出某种而非另一种选择。这种出现离散结果的情况是非常普遍的，广泛出现与各类统计应用中。林火会不会发生，最后只能将结果（因变量）分为“林火发生”和“林火未发生”两类定性的答案，而不可能存在介于发生和未发生之间的其他结果。我们研究林火预测预报关也的是与林火相关的地形地貌、土壤、植被、气象、社会等火险因子（自变量）对引发林火的贡献值有多大。但在使用机器学习进行回归运算时，因变量通常要转换成二值变量，即 1

或 0 进行运算，因此我们必须找到将贡献度转化为 1 到 0 之间概率并进一步转化为二值的方法。

$$\hat{p} = \sigma(\theta^T \cdot x_b) = \frac{1}{1 + e^{-\theta^T \cdot x_b}}$$

$$\hat{y} = \begin{cases} 1, & \hat{p} \geq 0.5 \\ 0, & \hat{p} \leq 0.5 \end{cases}$$



这里我们使用常见的线性回归转逻辑回归公式，其中， Θ 为因素权重， x 为自变量矩阵。我们会发现 p 函数正好在整个实数空间上分布在 0,1 之间，这符合我们对此函数的预期。

然后我们应确定森林起火与哪些因素有关，根据查询相关论文[1]，我们提出以下表格

诱火因子	自变量	变量类型
地形	海拔(m)	定量变量
	地貌	分类变量
	坡向	分类变量
	坡度(°)	定量变量
	坡位	分类变量
土壤	土壤类型	分类变量
	土壤质地	分类变量
	土壤厚度(m)	定量变量

植被	植被类型	分类变量
	植被总盖度(%)	定量变量
	灌木优势种	分类变量
	灌木盖度(%)	定量变量
	草本优势种	分类变量
	草本盖度(%)	定量变量
	优势树种	分类变量
	蓄积 (m ³)	定量变量
气象	相对湿度(%)	定量变量
	气温(°C)	定量变量
	降雨(雪)量(mm)	定量变量
	降雨雪后日数(d)	定量变量
	风速 (m/s)	定量变量
社会	客流量	定量变量
	客流流动速度	定量变量

考虑到可行性及论文中给出的相关因素权重,我们选取了其中的海拔、坡度、坡向、风速、温度、客流量以及使用 Erdas 对方山地区遥感图像求得植被系数 e 作为植物相关变量的替代作为我们的研究对象。

3.1.2 数据获取

通过相关论文我们可以得到房山区 1995-1999 年的详细火灾报告,因此我们选取房山区作为这次机器学习的样本地区,然后我们可以在中国知网上下载到房山区的 dem 和遥感森林指数分布并在中国气象数据网 (<http://data.cma.gov.cn/>) 上得到房山区的气候指标,通过 arcgis 我们可将这些原始数据依照其地理坐标进行整理和进一步处理,将大量信息捆绑为伴地理坐标的数据 batch,得出伴随地理拓扑关系的属性信息,以便于我们生成进一步机器学习所需的样本。(其中我们使用居民点的密集程度代指客流量)

3.1.3 基于机器学习的逻辑回归

在我们得到全面且大量的样本数据后，我们应注意，不同的起火因子对最终的火灾发生的贡献度并不一定符合线性回归，因此使用 matlab 求解单一回归的方法并不可行，因此我们使用 python 中 scikit-learn 包所支持的机器学习中的多项式回归进行逻辑回归。核心代码如下

```
1 from sklearn.preprocessing import PolynomialFeatures
2 from sklearn.pipeline import Pipeline
3 from sklearn.preprocessing import StandardScaler
4
5 def PolynomialLogisticRegression(degree):
6     return Pipeline([
7         ('poly', PolynomialFeatures(degree=degree)),
8         ('std_scaler', StandardScaler()),
9         ('log_reg', LogisticRegression())
10    ])

```

```
1 poly_log_reg2 = PolynomialLogisticRegression(degree=20)
2 poly_log_reg2.fit(X_train, y_train)

```

(其中，多项式维度为 20)

现在，我们就可以得到训练之后的起火因素的逻辑回归模型了。

3.1.4 模型预测

得到训练后的模型后，我们就可以针对上方山的观测数据预测上方山地区的火灾发生概率，但是注意，因为时间和人力的限制，我们所采用的为突变点算法，即找出测量自变量的突变点，并只对突变点进行模型预测，这样虽然得到的图并不平滑且有一定误差，但其一可以大大降低计算量，其二多项式回归所得到的函数一定具有任意阶导数，所以其函数一定在任意点连续，所以误差也一定在固定区间之内。

3.2 游客突发状况模型

3.2.1 问题分析和模型假设

虽然每一名游客的身体情况精神状态不同,但是,每一名游客上山后都会类似的原因导致出现意外情况(包括跌打损伤,突发心脏疾病,中暑等登山时常发生且与山体不同系数有关的突发状况),而影响这些原因的因素也是固定的,有限的,而每一名游客上山后都会且肯定会处于“出现突发状况”和“未出现突发状况”两种状态。因此,我们从长时间和大量游客的角度考虑,忽略时间和个体差异的影响,认为每一个游客在任何时候发生突发状态的概率只与与地理因素唯一相关的 p 值有关,那么,全山突发状态发生概率就符合 $\lambda = p$ 的泊松分布。

$$P(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}, k = 0, 1, \dots$$

$$\psi(t) = \exp\{\lambda(e^{it} - 1)\}$$

这样,我们只需要知道不同地区的 p 值,就可以使用泊松分布模型,求得在一定地区发生意外的概率和数学期望

3.2.2 p 值的求解

在这个问题中我们只关心不同区域意外状况概率的比值,以此来确定安全网的布控,但我们并不关心概率的真实值。因此,我们可以假设变量 w 为单位权,并认为 p 值除与不同诱发因子 a 有关外,只与自然因子 w 有挂,这样我们根据权重的传递公式:

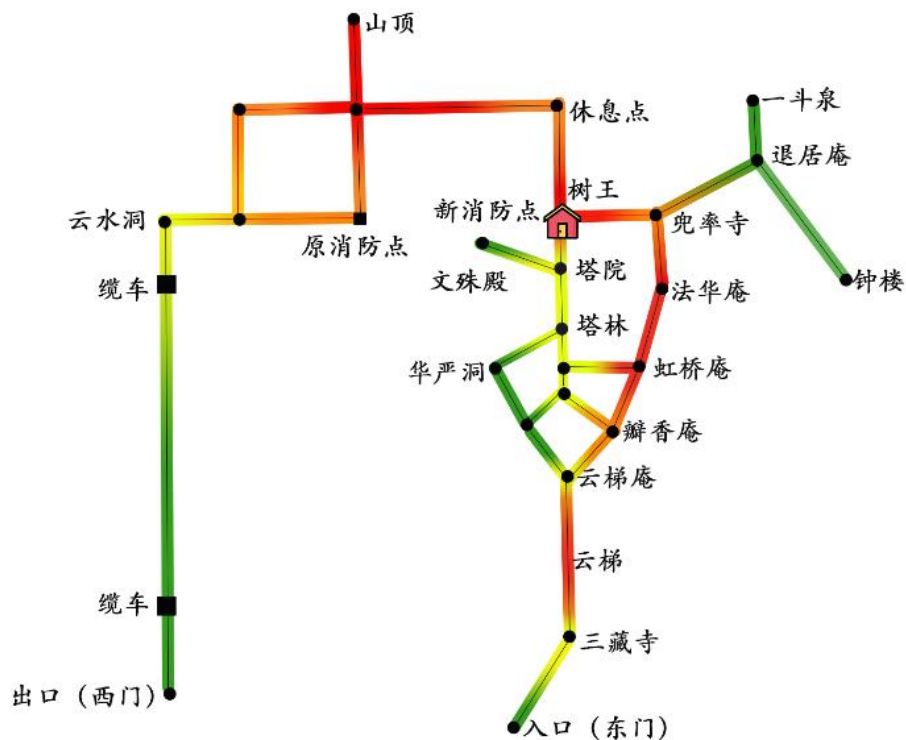
$$x'_i = (x_i - \frac{\sum_{i=1}^n x_i}{n}) / (\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2)$$

$$\frac{1}{w_p} = \sqrt{\sum_{i=1}^n (\frac{x'_i}{w})^2}$$

就可以得到每个因子 p 所对应的权重 w。

3.2.3 线路权值计算

我们认为游客并不会脱离已有路线行动,所以全部的意外状况也应全部发生在已有道路上。因此,我们仅对已有道路通过公式 () 进行权值根据实地测量值 (依然使用突变点计算) 进行计算,就可以得到对于游客突发状况的预警与处理的道路权重图



有效发现，在并在生成阶段就能控制火情，而如果在生成阶段未能识别火情，就要确保在单点蔓延阶段有足够功率的灭火设施，确定控制火情。

但是，要注意的是，不同阶段的火情也需要不同程度的防火器具来防治。而如果能早发现早灭火的话我们就不需要使用过于昂贵的高等级灭火器材，所以我们先通过相关网站查得不同的火灾防治器械，及其参考价格 c:

装置名	花费	覆盖范围	功用	火灾阶段
人	一份工资	30-50 米	发现火源、灭火器具	所有阶段的火灾
山火拍	20 元/把	3 米	灭火器具	诱发、生成阶段
风力灭火器	600 元	6 米	灭火器具	诱发、生成阶段
背负式水枪灭火器	400 元	20 米	灭火器具	全部阶段
红外探测器	2000 元	60 米	发现火灾	生成阶段
红外传感器	370 元	20 米	发现火灾	所有阶段

4.2.3 游客突发情况控制

其次，通过泰山山林安全组委会发布的 2018 年游客突发情况及解决情况统计表，与网上所查询得到的各类疾病控制情况如下表：

突发状况	严重程度(s)	可能造成损失(c)	安全时间	影响因素
擦伤	低	0-20	——	坡度，路况
扭伤	低	300-500	——	坡度，路况
骨折	中	6000-30000	6-8 小时	坡度，路况
中暑	中	0-10	2-3 小时	天气状况，身体状况，

心脏病	高	10000-20000	10 分钟	身体状况, 运动量
低血糖	中	0-10	——	身体状况, 运动量
痛风	高	0-2000	3 分钟	身体状况, 运动量
高血压	高	100-300	1 小时	身体状况, 运动量

因此, 我们本着以人身生命安全为最优先的方针, 我们应保证受伤人员在其安全时间内得到救治, 患烈性疾病的旅客也可以在安全事件内送往山下进行救治。所以, 无论旅客在山中何点发生了意外情况, 我们都应保证旅客可以正常与指挥中心去的联络, 并且指挥中心可以在规定时间内赶到事发地点并在必要时及时送往山下。但是要注意, 在山中有信号的部分我们可以使用手机报警, 而在没有信号的部分我们只能考虑安全电铃, 这时就还要考虑电铃的有效覆盖面积。

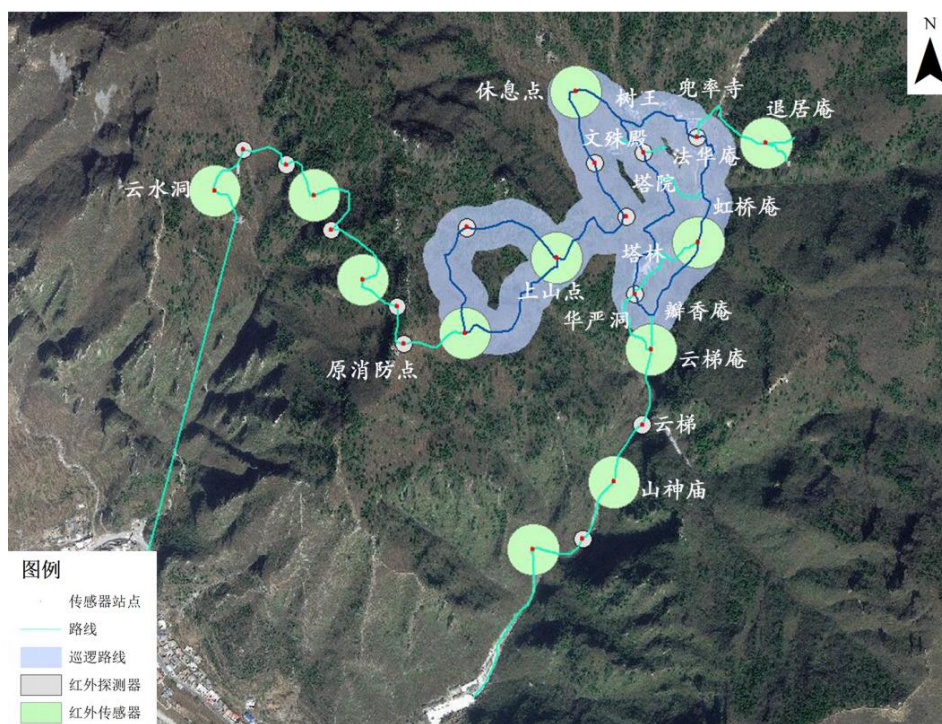
最后, 规范接下来可行性方案中所使用的的数据量和概念。

- 1、防火设施的有效范围以其公示的安全范围为准, 安全电铃的有效范围以游客的可视距离为准。
- 2、巡逻队默认在其巡逻路线的三个等分点上。
- 3、巡逻速度为 2m, 接警上山速度为 4 米, 接警下山速度为 6 米, 运送病人速度默认为 4m
- 4、在火灾的一二阶段默认任何方式均可灭火, 在火灾的第三阶段默认必须消防栓或消防水枪才可以灭火

4.3 安全控制网搭建方案

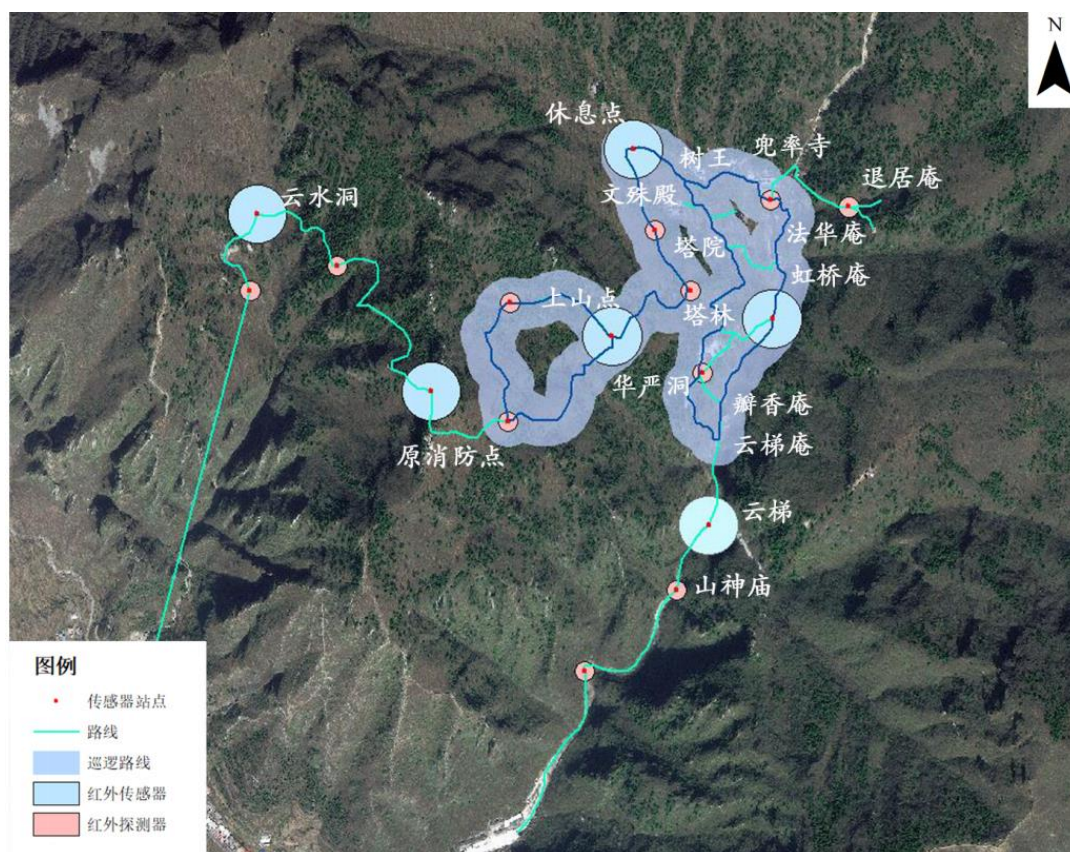
上方山的开发未必能得到消防队的重视,并派遣常驻人力上山,因此本次控制网方案分为有人方案和无人方案,而根据拓扑模型中重要点覆盖与否所决定的 i 的取值,巡逻队数量的最优解分为 2 和 3。所以,本次规划根据 N 的不同一共提出三个方案。

一、人机结合方案



人机结合方案将人工巡逻与传感器网络探测相结合,派遣两个小队沿巡逻路线,经过林火与人员危险高发地带,可对道路缓冲区 50 米左右范围缓冲区内的状况做出有限的监测和处理。而在危险系数较低区域布设两级传感器网络,第一级红外热探测器,第二级红外传感器,整套系统将人员火灾险情防控与传感器网全天候全时段监控结合,具有较高的安全系数和可靠性。

二、人工方案



人工方案采用三组巡逻队伍，增加一组巡逻队伍以减少传感器的数量，从而减少安全系统布设的总体费用，巡逻队伍的增加提高了沿途重要防火防人点的重返机会，提高巡逻路线沿途的安全系数，补充传感器减少造成的探测空缺，保证系统可靠性不出现较大变化。

三、智能方案



智能方案采用纯传感器网络监控危险状况，具有较高的智能性，高效性，可进行全天候，全时段监控，系统布设后具有较小的人力成本，但面对较复杂的林火状况，相对人工巡逻纯传感器安全网络鲁棒性较差。

4.4 方案的评价

根据之前的参数，我们可以使用 arcgis 中的缓冲区分析，得到不同安全装置的覆盖区域。因此，其一种缓冲区所不覆盖的区域就表示了一种安全隐患的可能发生区域，因此我们使用以下公式计算每套方案的安全系数 s ，和可能经济损失 c 。


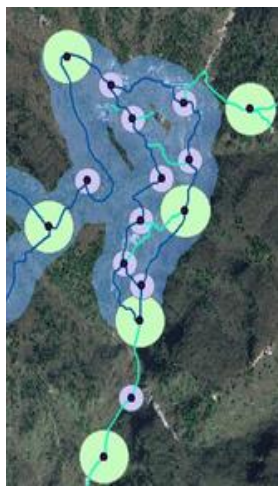

$$S = \sum_{i=1}^n \left(\frac{s_i}{S} \cdot \left| \frac{\lambda_{in} + \lambda_{out}}{\lambda_{in} - \lambda_{out}} \right| \cdot P_i \right)$$

$$C_{person} = \sum_{i=1}^{n-4} \left(\frac{S_i}{S} \cdot \left| \frac{\lambda_{in} + \lambda_{out}}{\lambda_{in} - \lambda_{out}} \right| \cdot P_i \cdot C_i \right)$$

$$C_{fire} = \sum_{i=1}^4 \left(\frac{S_i}{S} \cdot P_i \cdot C_i \right)$$

$$C = C_{fire} + C_{person}$$

然后我们可以先将 arcgis 中的数据库导入 SQL 数据库中, 用 SQL 数据库的计算器来计算出整个方案的花费, 安全系数, 可能经济损失 (客流量取现在的三倍), 展示如下

方案	智能操作	人机结合	人力
示意图			
方案评价	费用: 22500 元 安全系数: 7.79 可能经济损失: 36000 元/年	费用: 19250 元加两份工资 安全系数: 8.65 可能经济损失: 12000 元/年	费用: 12500 元加三分工资 安全系数: 7.62 可能经济损失: 32000 元/年

五、模型稳定性分析

本专题研究中，我们使用了图论模型与概率模型最终得到了安全控制网模型，这个过程本身都是在严格精度控制之下的。但是，我们要注意所有初始数据都是有其误差的。所以，我们要从最下层的模型进行逆推，以确定整个模型对误差变动的稳定性。

首先，对于图论模型，我们完全使用了简化后的拓扑图进行研究，而且，也使用了非最优化求解的遗传算法进行求解，因此，并没有误差可言。而对于概率模型，我们可以使用测量学中的误差传递公式，计算系统的中误差：

$$R - RMSE = \sqrt{\sum_{i=1}^n S_i^2 / n}$$

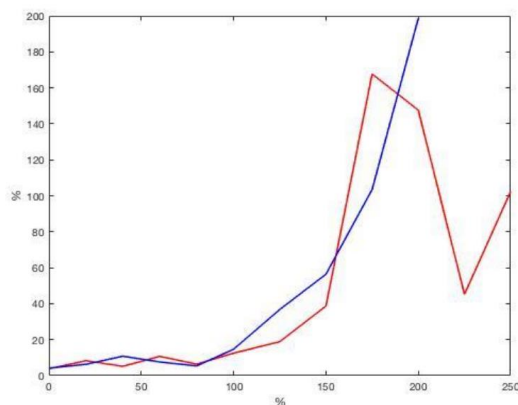
其中，R 为系统误差（仪器误差），RMSE 为单位权中误差，S 为每项方差。然后，我们列举全部模型中用到的测量变量，并对其做误差分析。

数据名称	标准化样本方差 (S)	单位权中误差 (RMSE)
海拔/m	0	0.5
遥感红波段	0	0.5
遥感绿波段	0	0.5
风速/(m/s)	0.2	0.325
温度/°C	0.05	0.0125
运动量/cal	0.05	0.25
客流量/人	0.5	1.25

之后我们可以发现除使用机器学习得到的逻辑回归模型外，其他模型均可用数学方法证明：其任意阶导数均可解，所以可以使用中误差传递公式，计算出最终中误差。

因此，我们先解决逻辑回归部分的误差分析，首先我们先解剖其中任意两个

神经节点, 通过在固定误差的一定百分比之内更改输入参数, 我们可以观察这两个神经元的变化情况



我们可以发现在控制误差的 100% 范围内, 神经节点所产生的变动可以控制在 10% 之内。但是超过 150% 的时候, 神经网络就会立刻失去稳定性。因此, 我们应该在实际测试时注意超限误差的产生。

而对于其他误差, 我们使用中误差的传递公式。

$$\frac{1}{m_l} = \sqrt{\sum_{i=1}^n \left(\frac{f'(x_i)}{m_i}\right)^2}$$

就可以得到影响最终测试方案的两个参数安全系数 s 和可能经济损失 c 的中误差。

	安全系数 (s)	可能经济损失 (c)
误差分析	± 0.07	± 268.74 元

参考文献:

【1】房山市 1995-1999 年火灾统计及其解决情况

房山市气象局

- | | |
|----------------------------|-----|
| 【2】北京房山林火发生预测模型及小班火险等级区划研究 | 林贺 |
| 【3】全局最优警车巡逻区域最大覆盖调度策略 | 吴思远 |
| 【4】基于信息扩散理论的区域森林火灾风险预测研究 | 刘佳奇 |