

基于二型模糊逻辑的交通流量预测

张伟斌, 胡怀中, 刘文江

(西安交通大学电子与信息工程学院, 710049, 西安)

摘要: 提出了一种改进的模糊 c 均值聚类算法, 该算法将模糊聚类的对象从单值扩展到区间, 在构造二型模糊系统时, 通过对历史数据的学习提取二型模糊规则, 克服了专家方法不能对未知领域提取规则的不足. 在此基础上, 针对智能交通系统, 提出一种新的基于二型模糊逻辑的交通流量预测方法. 该方法应用区间型二型模糊集具有上下限隶属度函数的性质构造预测区间, 适合于处理具有复杂不确定性的情况. 通过隶属度函数可以反映出该区间中预测值的可靠性, 从而克服了其他预测方法仅给出单值且稳定性不高的缺点. 仿真结果表明, 基于二型模糊逻辑的流量预测区间具有较高的准确度, 其平均相对误差低于 6%.

关键词: 流量预测; 二型模糊逻辑; 模糊聚类

中图分类号: TP 273.4 **文献标识码:** A **文章编号:** 0253-987X(2007)10-1160-05

Traffic Flow Forecast Based on Type-2 Fuzzy Logic Approach

Zhang Weibin, Hu Huaizhong, Liu Wenjiang

(School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: An improved fuzzy c -means clustering algorithm is proposed, in which the fuzzy clustering object is extended from single value to the interval data. In constructing the type-2 fuzzy system, the type-2 fuzzy rules are extracted from learning history data and the deficiency that human expert method can not extract rules in unknown systems is avoided. Based on the proposed algorithm, a novel traffic forecasting method based on type-2 fuzzy logic is proposed for intelligent traffic systems, in which the property that the interval type-2 fuzzy logic set possesses membership functions with upper and lower limit is utilized to create forecasting interval, which are suitable for handling the situations with complicated uncertainties. The reliability of the predicted values in that interval can be reflected through the membership function, thus the limits produced by other forecasting methods that only single value is given and are lack of stability is overcome. Simulation results show that the proposed algorithm has high accuracy and the average relative error is less than 6%.

Keywords: traffic flow forecasting; type-2 fuzzy logic; fuzzy clustering

交通流量预测是智能交通系统的常见组成部分, 对于交通控制策略的设计和应用有着重要的意义. 智能交通系统的性能在很大程度上依赖于交通流量预测的质量. 近年来有不少关于交通流量预测的研究, 比如 Turochy 使用非参数回归的预测方法^[1], Shiliang Sun 应用贝叶斯网络的预测方法^[2]

等, 但是这些方法的预测结果皆为单一数值, 难以确定预测的可靠性^[3]. 本文将二型模糊逻辑用于对交通流量的预测, 得到了流量的预测区间, 从而能够更加准确地给出预测流量的范围和分布.

二型模糊集是一型模糊集的扩展, 适于处理模糊程度复杂的系统^[4], 比如人说话时的不同口音,

信号处理的时变系统等,已经获得一些成功应用. 区间型二型模糊集是二型模糊集的简化,有效降低了二型模糊系统繁琐的计算量,但是如何从观测数据中提取模糊规则仍然是系统建模的关键. 本文采用一种改进的模糊 c 均值聚类算法从区间型二型模糊集中提取模糊规则,建立模糊系统,从而实现对交通流量区间的准确预测.

1 二型模糊理论简介

1.1 二型模糊系统

模糊系统基于模糊集合来构造,并运用相应的模糊逻辑推理和精确化来实现特定的系统功能,但在处理实际对象的不确定性时,传统一型模糊集具有一定的局限性. 将一型模糊集合扩展开来,进一步给出集合中隶属度的模糊程度,从而使描述集合模糊性的能力增强,这种扩展的模糊集合称为二型模糊集. 基于二型模糊集建立的系统称为二型模糊系统,适合处理时变特征难以用数学语言描述的系统^[4]. 二型模糊系统的基础是二型模糊集合理论^[5-6].

定义 1 如果 $\tilde{A}$ 是定义域在 X 上的二型模糊集, A 为 X 上的一型模糊集, $x \in X$, 则有

$$\tilde{A} = \{((x, u_A), \mu_{\tilde{A}}(x, u_A)) \mid \forall x \in X, \forall u_A \in J_x \subseteq [0, 1]\} \quad (1)$$

式中: u_A 为一型模糊集隶属度函数; $\mu_{\tilde{A}}$ 为二型模糊集隶属度函数. 二型模糊集概念如图 1 所示.

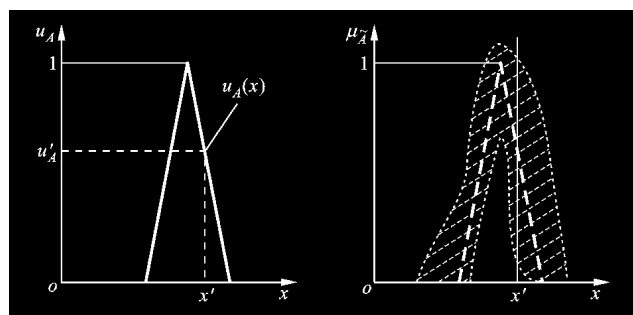


图 1 二型模糊集概念

将图 1a 中的一型隶属度函数值模糊化,使三角形的取值向左或者向右移动,移动的距离不定,结果如图 1b 所示. 对于数值 x 的某个具体的值 x' ,对应的函数值不再是某个单值 u'_A ,而是与直线 $x=x'$ 相交的一个取值范围. 对应不同的分布,这个取值范围内的值具有不同的可能性. 令所有的 $x \in X$ 模糊化,就得到了二型模糊集.

二型模糊系统的构造与一型模糊系统相似,如图 2 所示,一般包括输入模糊器、规则库、推理引擎、

降型器和精确器.

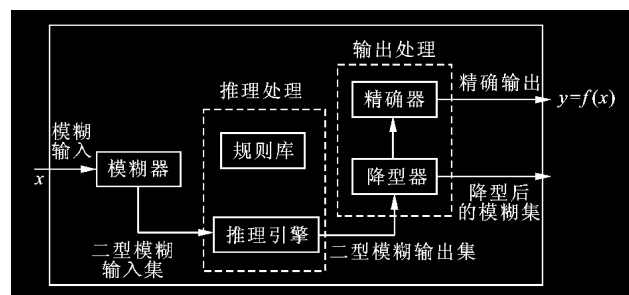


图 2 二型模糊系统

1.2 区间型二型模糊集

二型模糊集的运算非常复杂,尤其是降型运算. 为了简化运算,设次隶属度为 1,这样二型模糊集 $\tilde{A}$ 就成为区间型二型模糊集,定义如下.

定义 2 $\tilde{A}$ 称为区间型二型模糊集,当 $\forall x \in X$, $\mu_{\tilde{A}}(x, u_A) = 1$ 时,有

$$\tilde{A} = \int_{x \in X} \int_{u_A \in J_x} 1/(x, u_A) du_A dx \quad J_x \subseteq [0, 1] \quad (2)$$

2 二型模糊逻辑与交通流量预测

2.1 二型模糊集隶属度函数

区间型二型模糊集具有上限和下限隶属度函数,这个性质可以用来产生交通流量的预测区间. 本文应用 Gauss 型的主隶属度函数,如图 3 所示,图中 m 为期望.

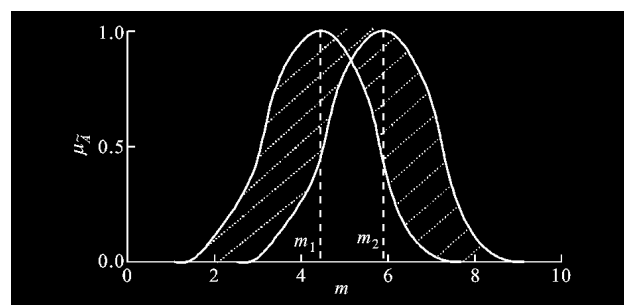


图 3 区间二型模糊集隶属度函数示意图

设标准差 σ 不变,期望 m 不确定,则有

$$\mu_{\tilde{A}}(x) = \exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right] \quad m \in [m_1, m_2] \quad (3)$$

对于不同的 m 值,都对应一条相应的隶属度曲线. m_1 、 m_2 的选择基于历史数据. 在区间型二型模糊集中,上限隶属度函数定义为

$$\bar{\mu}_{\tilde{A}}(x) = \begin{cases} N(m_1, \sigma; x) & x \leq m_1 \\ 1 & m_1 < x \leq m_2 \\ N(m_2, \sigma; x) & x > m_2 \end{cases} \quad (4)$$

下限隶属度函数定义为

$$\underline{\mu}_A(x) = \begin{cases} N(m_2, \sigma; x) & x \leq (m_1 + m_2)/2 \\ N(m_1, \sigma; x) & x \geq (m_1 + m_2)/2 \end{cases} \quad (5)$$

2.2 二型模糊预测

假设观测同一地区各星期的交通流量. 如果观测的时间区间很短, 例如 10 min, 那么该地区每星期中各天的交通流量是随机的, 它受到各种因素的影响, 如天气, 交通突发事件等, 但是如果观测区间增大到相当的范围, 那么该地区每天的交通模式在某个范围内则是重复的^[3]. 本文应用二型模糊集合的隶属度函数来模拟这个重复的范围, 实现流量预测. 二型模糊系统设计如下.

2.2.1 模糊器设计 如上节所述, 二型模糊集的上限隶属度函数和下限隶属度函数可以用来产生预测值区间. 本文应用式(3)预测交通流量. 设 σ 固定, 观测区间 $T_k = [t_{k1}, t_{k2}]$, p_{ki} 为 T_k 内第 i 天的峰值. 令 $m_1 = \min(p_{ki})$, $m_2 = \max(p_{ki})$, $\Delta m = (m_2 - m_1)/n$, 则 m 的次隶属度函数为

$$\mu_{mi} = \mu_m(m_i, m_{i+1}) = f(m_i, m_{i+1})/f(m_1, m_2) \quad (6)$$

式中: $f(m_i, m_{i+1})$ 为历史数据中峰值 $p_k \in [m_i, m_{i+1}]$ 的频率.

2.2.2 构造模糊规则 模糊规则用来处理隶属度函数产生的模糊输入. 设 x_i 是输入的第 i 个分量, F_{li} 是第 l 条规则的第 i 个前件分量, 是与 x_i 对应的模糊集, y_l 表示第 l 条规则的输出函数. 本文使用 TS 型的规则后件, 模糊规则如下

$$\begin{aligned} &\text{if } x_1 \text{ is } F_{11} \text{ and } x_2 \text{ is } F_{12} \cdots x_p \text{ is } F_{1p} \\ &\text{then } y_1 = f_1(x_1, x_2, \dots, x_p) \end{aligned}$$

本文应用模糊聚类算法提取模糊规则.

2.2.3 推理引擎 模糊推理是模糊系统的关键, 推理引擎根据模糊输入和规则前件来计算二型输出隶属度函数

$$\mu_{\hat{B}}(y) = \prod_{l=1}^N \mu_{B_l}(y) \quad (7)$$

式中: $\hat{B}$ 为二型输出模糊集; $\mu_{B_l}(y) = \prod F_{li}(x_i)$, x_i 为输入数据.

2.2.4 降型器 降型是二型模糊系统的特点. 推理引擎输出的结果是二型模糊集, 在精确化之前, 必须降型. 本文采用质心降型方法^[6]得到质心区间 $Y_c(x)$.

2.2.5 精确器 精确化是取得最终预测结果的最后一步. 本文使用质心法, 精确器的输出结果为

$$y_c(x) = \frac{\sum_{i=1}^N y_i \mu_{\hat{B}}(y_i)}{\sum_{i=1}^N \mu_{\hat{B}}(y_i)} \quad (8)$$

精确器的输出结果即为具体的预测区间.

2.3 二型模糊规则提取

2.3.1 改进的模糊 c 均值聚类算法 模糊聚类算法是提取模糊规则的有效方法. 其中, 模糊 c 均值聚类算法 (FCM) 是目前应用最广泛的聚类算法之一^[7], 但其处理对象仅限于点集. 对于特殊类型的数据, 如区间型数据, FCM 算法无法直接处理. 本文扩展了样本方差的概念, 使之从点集扩充到区间集, 并基于此概念, 建立新的泛函对样本数据类型为区间的数据集进行聚类划分.

设数据集的样本元素为区间 $x_k = [x_{k1}, x_{k2}]$, 第 i 个聚类中心为 $[p_{i1}, p_{i2}]$, 则样本 x_k 到聚类原型 p_i 的距离定义为

$$d_{ik} = (\|x_{k1} - p_{i1}\|^2 + \|x_{k2} - p_{i2}\|^2)^{1/2} \quad (9)$$

设 U 为隶属度函数集, P 为聚类原型集, 令

$$\begin{aligned} J'_\omega(U, P) &= \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^\omega (d_{ik})^2 = \\ &\sum_{k=1}^n \sum_{i=1}^c (u_{ik})^\omega (\|x_{k1} - p_{i1}\|^2 + \|x_{k2} - p_{i2}\|^2) \\ &1 \leq \omega \leq \infty \end{aligned} \quad (10)$$

式中: u_{ik} 为隶属度函数; ω 为加权指数; c 为聚类类别数; k 为样本序号; n 为样本数量.

聚类的准则为求取 $J'_\omega(U, P)$ 的极小值 $\min\{J'_\omega(U, P)\}$, 用拉格朗日乘数法求解得

$$F = \sum_{i=1}^c (u_{ik})^\omega (d_{ik})^2 + \lambda (\sum_{i=1}^c u_{ik} - 1) \quad (11)$$

分别取 $\frac{\partial F}{\partial \lambda} = 0$, $\frac{\partial F}{\partial u_{ik}} = 0$, 解得

$$\begin{aligned} u_{ik} &= 1 / \left\{ \sum_{l=1}^c \left[\frac{d_{lk}^2}{d_{ik}^2} \right]^{\frac{2}{\omega-1}} \right\} = \\ &1 / \left\{ \sum_{l=1}^c \left[\frac{(\|x_{k1} - p_{l1}\|^2 + \|x_{k2} - p_{l2}\|^2)^{1/2}}{(\|x_{k1} - p_{i1}\|^2 + \|x_{k2} - p_{i2}\|^2)^{1/2}} \right]^{\frac{2}{\omega-1}} \right\} \end{aligned} \quad (12)$$

类似地, 分别令 $\frac{\partial F}{\partial p_{i1}} = 0$, $\frac{\partial F}{\partial p_{i2}} = 0$, 解得

$$p_{i1} = \frac{1}{\sum_{k=1}^n (u_{ik})^\omega} \sum_{k=1}^n (u_{ik})^\omega x_{k1} \quad (13)$$

$$p_{i2} = \frac{1}{\sum_{k=1}^n (u_{ik})^\omega} \sum_{k=1}^n (u_{ik})^\omega x_{k2} \quad (14)$$

这样,根据数据集 $X = \{(x_{k1}, x_{k2})\}$ 、聚类类别数 c 和加权指数 ω ,就可确定模糊分类矩阵 $[u_{jk}]$ 和聚类区间中心 $[p_{i1}, p_{i2}]$.

算法的具体实现采用迭代方法. 给定聚类类别数 $c, 2 \leq c \leq n, n$ 为样本数据个数. 设定迭代停止阈值 ϵ, b 为迭代计数器,迭代终止条件为

$$\|P_{jk}^{(b)} - P_{jk}^{(b+1)}\| < \epsilon \quad (15)$$

迭代计算

$$p_i^{(b+1)} = \frac{\sum_{k=1}^n (\mu_{ik}^{(b+1)})^\omega x_k}{\sum_{k=1}^n (\mu_{ik}^{(b+1)})^\omega} \quad i = 1, 2, \dots, c \quad (16)$$

即可确定聚类中心 $[p_{i1}, p_{i2}]$ 和模糊分类矩阵 $[u_{jk}]$.

2.3.2 聚类优化 模糊聚类产生的聚类结果往往包含了一些多余的分类,因此有必要对聚类结果进行优化. 本文所采用的优化方法是首先给出聚类数量估计值 c 的上限,然后评估聚类的相似性,对相似性超过阈值 $\alpha \in [0, 1]$ 的类进行合并. 定义在 X 上的模糊聚类的相似性定义^[8]为

$$S_{ij} = \frac{\sum_{k=1}^N \min(u_{ik}, u_{jk})}{\min(\sum_{k=1}^N u_{ik}, \sum_{k=1}^N u_{jk})} \quad (17)$$

根据每个聚类产生的分类可以建立一条相应的 IF-THEN 模糊规则.

3 验证与仿真

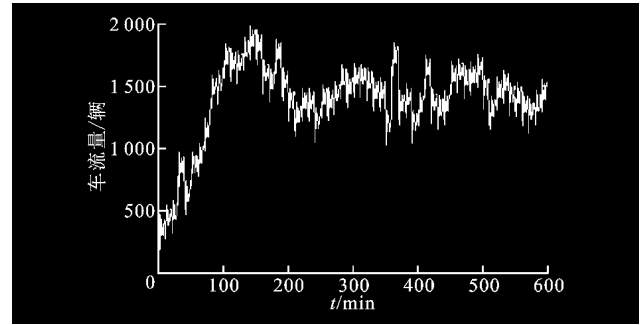
本文实验采用 Matlab7 编制仿真程序,验证数据取自美国加州 PATH 资料库,收集于 Nimitz 高速公路的一段 11.26 km(7 英里)长的路段^[9]. 图 4 为连续 5 周中每星期二交通流量随时间的变化情况.

从图 4 的数据可以看出,尽管每个时刻的交通流量是不断变化的,但是不同日期的固定时段的流量变化趋势却是类似的. 取每 10 min 作为时间间隔,应用本文模型,将前 4 周的数据作为历史数据,得到如图 5 所示的预测结果,图 5 中实际流量数据为图 4e 中数据取每 10 min 的平均值.

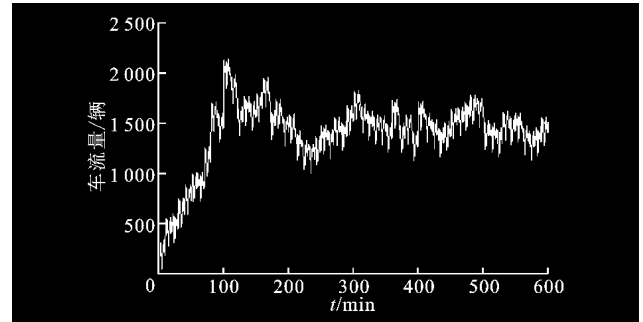
为检验预测算法的稳定性,共进行了 10 次流量预测,对预测结果使用平均相对误差 E_{MR} 进行评估

$$E_{MR} = \frac{\sum_{n=1}^N \sum_{i=1}^I |\hat{x}_i^n - x_i^n| / x_i^n}{NI} \quad (18)$$

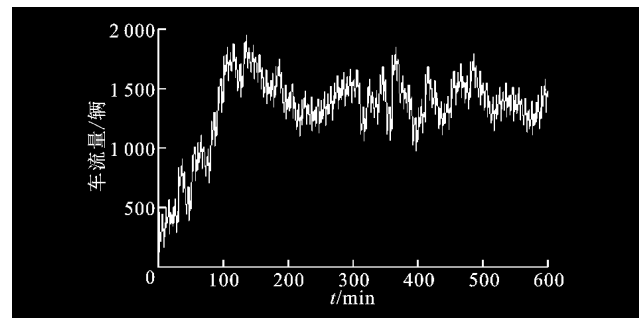
式中: $\hat{x}_i^{(n)}$ 为第 n 次预测中第 i 个时间段的流量预测中值,即预测流量上限与下限的平均值; N 为预测总数; I 为每次预测中包含的数据量.



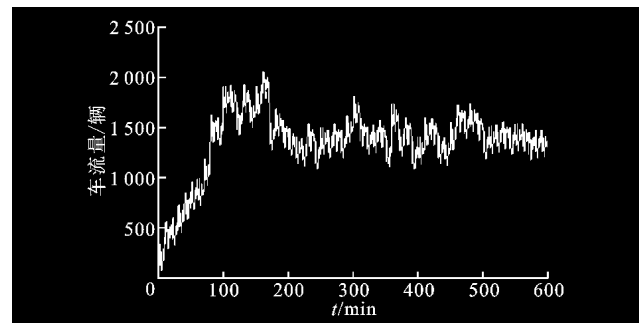
(a) 第1周



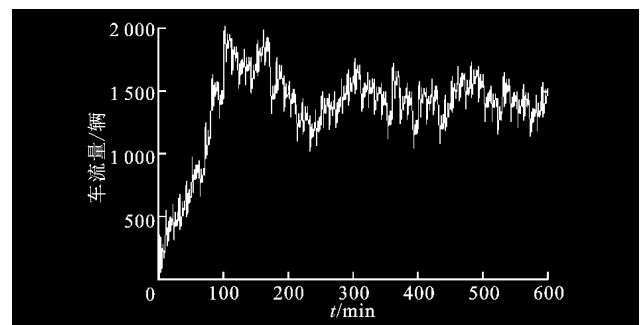
(b) 第2周



(c) 第3周



(d) 第4周



(e) 第5周

图4 某高速公路交通流量数据

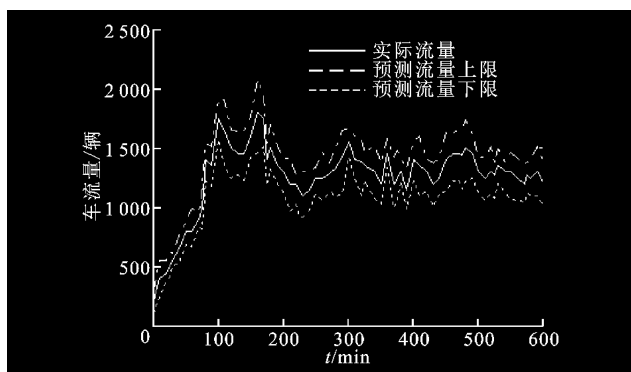


图5 预测流量与实际流量比较图

由式(18)计算得预测中值与时间值之间的平均相对误差为 5.78%。根据图 5 的预测结果以及式(18)的计算结果,可以得出预测区间与实际值比较吻合的结论。

4 结束语

本文基于二型模糊系统建立了交通流量预测模型,得到了交通流量的预测区间,使流量预测更加准确合理。在二型模糊系统的建模过程中,本文对广泛使用的模糊 c 均值聚类方法进行了改进,将处理对象从点集扩展到区间集,从而可以从区间型二型模糊系统中抽取模糊规则。仿真结果表明,本文的交通流量的区间预测方法具有较高的精确度。下一步的研究工作将着重于流量预测方法在智能交通系统中的应用。

参考文献:

- [1] Turochy R E, Pierce B D. Relating short-term traffic forecasting to current system state using nonparametric regression [C] // Proceedings of IEEE Intelligent Transportation Systems Conf. Piscataway, USA: IEEE, 2004: 239-244.
- [2] Shiliang S, Changshui Z, Guoqiang Y. A Bayesian network approach to traffic flow forecasting [J]. IEEE Transactions on Intelligent Transportation Systems, 2006, 7(1): 124-132.
- [3] Li L, Lin W H, Liu H. Type-2 fuzzy logic approach for short-term traffic forecasting [J]. IEE Proceeding of Intelligent Transportation System, 2006, 153(1): 33-40.
- [4] 陈薇, 孙增圻. 二型模糊系统研究与应用[J]. 模糊系统与数学, 2005, 19(1): 126-135.
Chen Wei, Sun Zengqi. Research on of type-2 fuzzy logic system and its application [J]. Fuzzy Systems and Mathematics, 2005, 19(1): 126-135.
- [5] Jerry M M, Robert I J, Feilong L. Interval type-2 fuzzy logic systems made simple [J]. IEEE Transactions on Fuzzy Systems, 2006, 14(6): 808-821.
- [6] Qilian L, Jerry M M. Interval type-2 fuzzy logic systems: theory and design [J]. IEEE Transactions on Fuzzy Systems, 2000, 8(5): 535-550.
- [7] Magne S. Supervised fuzzy clustering for rule extraction [J]. IEEE Transactions on Fuzzy Systems, 2000, 8(4): 416-424.
- [8] Uzey K, Magne S. Fuzzy clustering with volume prototypes and adaptive cluster merging [J]. IEEE Transactions on Fuzzy Systems, 2002, 10(6): 705-712.
- [9] Skabardonis A, Noeimi H, Petty K. Free way service patrol evaluation, UCB-ITS-PRR-94 [R]. Berkeley, USA: Institute of Transportation Studies, University of California at Berkeley. 1994.

(编辑 刘杨)