

## 基于全球 AIS 的多源航迹关联数据集使用说明

多源航迹关联数据集（Multi-source Track Association Dataset, MTAD）是依托国家自然科学基金、中国科协青年人才托举工程、山东省泰山学者人才工程，以全球 AIS 航迹数据为基础，由全球 AIS 航迹数据经栅格划分、自动中断和噪声添加处理步骤构建的，旨在解决智能关联算法研究对数据的迫切需求和多雷达协同观测航迹数据获取困难，填补航迹关联数据集的空白。该数据集目前已被用作海军“金海豚”杯竞赛科目 3 “大差异海上目标航迹关联”的专用数据集。

对全球 AIS 航迹数据进行栅格划分得到的全球栅格是 MTAD 数据集构建的基础，数据集集中的所有场景均从全球栅格中抽取产生。在全球地图中根据所有栅格中的 MMSI 数量绘制热力图，对 MMSI 数量进行可视化，结果如图 1 所示。从图 1 中可以看出，在重要港口地带，MMSI 数量较多，在远海区域，MMSI 数量较少。各航迹在全球各个海域均广泛分布，为数据集的构建提供了丰富的航迹资源。

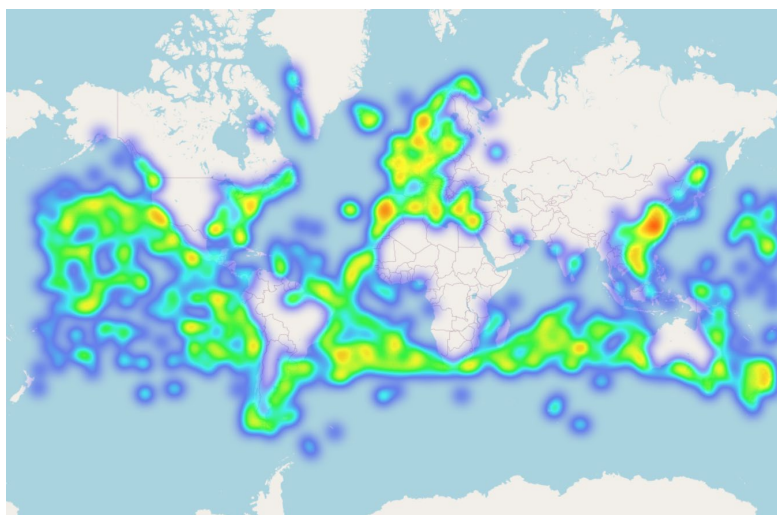


图 1 MMSI 数量热力图

MTAD 数据集结构如图 2 所示，包括训练集和测试集两大部分，训练集和测试集均由航迹信息表和关联映射表组成，均为 csv 格式文件，根据场景名称进行区分。该数据集共有航迹百万余条，其中训练集包含 5000 个场景样本，测试集包含 1000 个场景样本，每一个场景样本由几个到几百个数量不等的航迹构成，涵盖多种运动模式、多种目标类型和长度不等的持续时间。

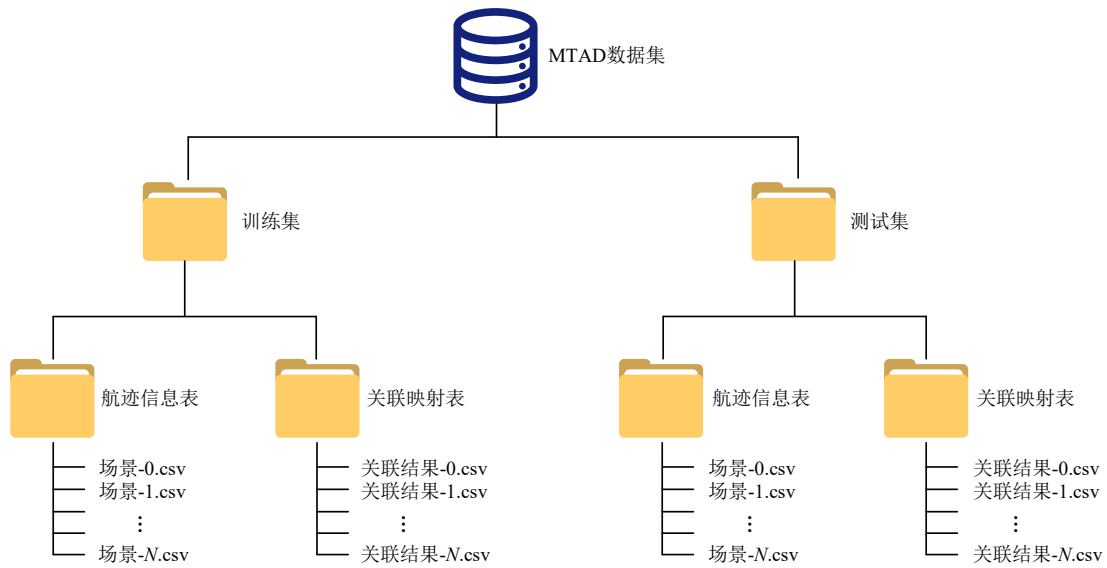
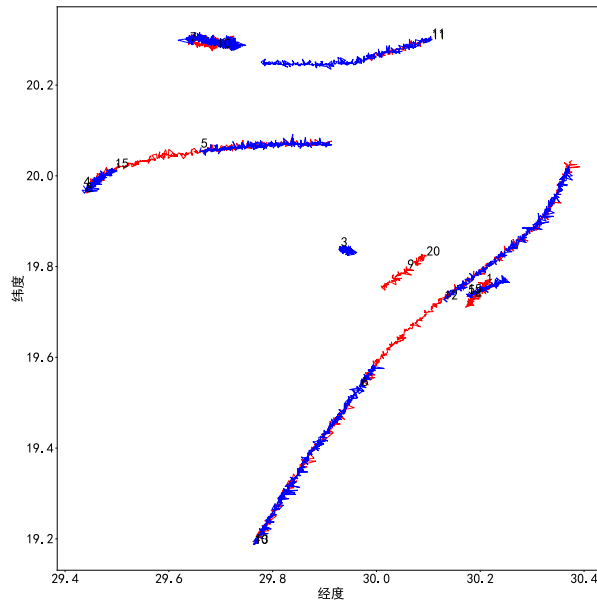


图 2 MTAD 数据集结构

MTAD 中的某个场景如图 3 所示，其中，图 3 (a) 中红色航迹表示信源一，蓝色航迹表示信源二，从图中可以清楚的看出航迹的多源现象和中断现象。图 3 (b) 对应的航迹信息表的属性包括{批号、信源号 (9001、9002，随机设置)、时间 (一天内的绝对秒)、经度 (度)、纬度 (度)、航速 (米/秒)、航向 (度)}，在 csv 文件中的对应表示为{batch、source、time、lat、lon、vel、cou}。图 3 (c) 对应的关联映射表的属性包括{开始时间-结束时间-真值批号-信源号-航迹批号}，在 csv 文件中的对应表示为{t\_s、t\_e、mmsi、source、batch}。

多源航迹关联任务可以表述为：利用航迹信息表中的航迹信息，实现对于同一目标、不同信源之间航迹批号的统一。其中同一目标定义为在关联映射表中的 mmsi 号相同。例如，在关联映射表中，信源号为 9002、航迹批号为 15 的航迹和信源号为 9001、航迹批号为 3 的航迹来自不同的信源，但具有相同的 mmsi 号，因此两者为同一目标，是由于多源观测导致的航迹冗余，应实现其航迹批号的统一。

中断航迹接续关联任务可以表述为：利用航迹信息表中的航迹信息，实现对于同一目标、同一信源之间航迹批号的统一。其中同一目标定义为在关联映射表中的 mmsi 号除了最后一位之外均相同，mmsi 号的最后一位表示航迹中断的片段索引。例如，在关联映射表中，信源号为 9001、航迹批号为 9 的航迹和信源号为 9001、航迹批号为 20 的航迹来自同一信源且批号不同，但其 mmsi 号除了最后一位之外均相同，因此两者为同一目标，是由于航迹中断导致的航迹冗余，应实现其航迹批号的统一。



(a) 航迹图像

batch	source	time	lat	lon	vel	cou	t_s	t_e	mmsi	source	batch
15	9002	50	19.74077	30.17687	132.9237	194.9478	50	1499	413484880-3-0	9002	15
5	9001	50	19.73915	30.17706	60.2877	177.514	50	1509	413484880-3-0	9001	5
3	9002	51	19.84378	29.93119	151.9687	193.8181	51	1500	413521560-0-0	9002	3
15	9002	60.062	19.73479	30.17517	24.32831	79.4719	131	4877	461564169000-3-0	9001	13
3	9002	61.062	19.83692	29.9294	391.9288	76.25155	143	748	044219028351-5-0	9001	20
15	9002	70.125	19.735	30.17635	180.9622	112.4443	321	6414	998564169000-3-0	9002	10
5	9001	70.264	19.73351	30.17732	45.66069	16.8822	344	4006	032257021000-1-0	9002	11
3	9002	71.125	19.84125	29.94822	334.884	227.4482	527	2390	841257983600-0-0	9001	8
15	9002	80.188	19.73179	30.18462	275.5631	305.7589	560	1343	921574151045-0-0	9001	1
3	9002	81.188	19.83072	29.93602	242.931	312.8909	699	2300	877257021000-1-0	9001	11
15	9002	90.25	19.73928	30.17357	147.9888	128.2613	913	3035	072257983600-0-0	9002	4
5	9001	90.528	19.7376	30.17864	9.758672	193.7686	1070.735	2422	219028351-5-1	9001	9
3	9002	91.25	19.83841	29.92722	186.1404	73.50629	2731.543	9746	257983600-0-1	9001	15
15	9002	100.312	19.73501	30.17931	37.85973	15.45626	2811.259	4359	574151045-0-1	9001	4
3	9002	101.312	19.84087	29.93605	169.3418	196.785	5578.415	14090	564169000-3-1	9001	6
15	9002	110.375	19.73671	30.17981	127.3727	148.3504	6168.13	10132	257983600-0-1	9002	5
5	9001	110.792	19.73671	30.17841	47.92256	233.8468	6285.544	16918	257021000-1-1	9001	10
3	9002	111.375	19.83332	29.93363	54.61795	32.29776	7688.075	16563	257021000-1-1	9002	7
15	9002	120.438	19.73167	30.18311	76.85548	318.9951	8976.678	14040	564169000-3-1	9002	12

(b) 航迹信息表 (c) 关联映射表

图 3 MTAD 数据集场景示例

基于全球 AIS 的多源航迹关联数据集 MTAD，由海军航空大学制作，《电子与信息学报》编辑部具有编辑出版权等。

读者可免费使用该数据进行教学、科研等，但需在论文、报告等成果中引用或致谢。该数据禁止私自用于商业目的，如有商业需求，请与《电子与信息学报》编辑部联系。

若在论文、报告中使用该数据集，请引用以下文献：

崔亚奇，徐平亮，龚诚等. 基于全球 AIS 的多源航迹关联数据集[J]. 电子与信息学报，待出版，doi:10.11999/JEIT221202.