

面向情报感知的 多源多模态数据融合方法研究

白如江 鞠孜涵 张玉洁 张亚辉 冯梦莹

【摘要】[研究目的]针对多源多模态数据在情报感知应用中存在的问题,在梳理多源多模态数据融合方法的基础上提出一种新的多模态数据融合思路,以支撑细粒度情报感知的相关研究。[研究方法]首先,系统梳理了情报感知和多源多模态数据的特征,总结了多源多模态数据融合面临的挑战。其次,从不同的融合粒度详细调研了多源异构数据的融合方法;从不同的融合深度总结了多模态数据的语义融合方法。最后,提出了时序数据和地理空间数据辅助的多模态数据融合思路。[研究结论]面向情报感知的多源多模态数据融合中,细粒度的融合是基础,语义融合是核心,基于深度学习的融合方法是大势所趋,基于异质图神经网络的融合方法推动了情报感知更准确、全面和智慧的发展。

【关键词】情报感知;多源多模态数据;数据融合;细粒度;语义融合

【作者简介】白如江(通信作者)(1979-),男,博士,博士生导师,山东理工大学信息管理研究院,研究方向:科技情报分析、情报感知;鞠孜涵(1999-),女,硕士研究生,山东理工大学信息管理研究院,研究方向:科技情报分析、情报感知;张玉洁(1997-),男,硕士研究生,山东理工大学信息管理研究院,研究方向:科技情报分析;张亚辉(1997-),女,硕士研究生,山东理工大学信息管理研究院,研究方向:科技情报分析;冯梦莹(1999-),女,硕士研究生,山东理工大学信息管理研究院,研究方向:科技情报分析(淄博 255049)。

【原文出处】《情报杂志》(西安),2023.10.124~131

【基金项目】国家社会科学基金项目“多源数据融合驱动的智慧情报感知研究”(编号:21BTQ071)研究成果。

随着“互联网+”“5G”时代的到来,情报工作拓展到了科技、经济、社会等诸多领域,情报分析的数据源呈现出海量、多源、异构、多模态等特征,数据之间的关联更为隐蔽多变。但情报工作的本质始终是解决决策管理过程中信息不完备的问题^[1],这也使得情报感知成了新时代情报工作的核心。“情报感知”是对“未知”的问题进行感知探索,“早醒远眺”,对不确定的未来揭示预警,从而减少“意外”的发生^[2]。其中,“感”是指情报人员对数据的获取和处理;“知”则强调了数据的理解、评析与展望。由此可知,情报感知对数据优势的依赖性越来越强,两者密不可分。出色的情报感知工作必须要有丰富的数据支撑,进而发现数

据之间复杂的关联,挖掘出更深层、未知的研究内容,解决信息不完备情况下的决策问题。

大数据时代的数据往往呈现出碎片化、线索化的特征,无法提供完整的情报地图。因此,情报感知工作不仅需要多源多模态的数据支持,还需要将这些数据进行有效融合,形成格式统一、面向感知应用的数据集^[3]。只有通过数据之间的相互支持、补充和校对,才能提供更加准确、全面和及时的信息支撑情报感知工作。在此基础上,数据融合粒度的问题也应运而生。由于数据外部特征融合方法难免会带来宏观、粗糙、客观性差、层次浅的问题^[4],如何将数据的外部结构特征和内部语义特征深度融合,更加全面、深入、客

观的分析情报元素,支撑情报感知研究迫在眉睫。

1 相关研究

情报感知是对“未知”问题的进行感知和探索。早在2008年杜元清研究员提出“有组织的‘惦记’和‘监测’”^[5],就昭示着情报学者对于情报感知的认知追求。2016年杨峰等^[6]在“情景嵌入的突发事件情报感知”研究中,虽然没有给出“情报感知”的明确定义,但指出了要在多源异构细粒度情报资源的融合层面上讨论感知问题。2018年王延飞教授^[7]首次界定了“情报感知”的概念,并认为情报工作必须重视对情报需求、情报对象和情报任务的感知和刻画。除此之外,杨峰等^[8]探讨了基于情景相似度的突发事件情报感知方法,并指出全源情报资料是感知的基础。马红岩等^[9]从科技情报感知出发,构建了多源信息融合的新模式。白如江等^[10]提出了智慧情报感知的内涵,将因果推断、事件溯因等方法引入情报感知研究,并指出了语义级多源数据融合的重要性。在上述情报感知研究中,均在不同程度上指出了数据融合在情报感知中的重要性,学界对技术赋能并融合多源多模态的数据进行情报感知研究已经形成共识。

数据融合(Data Fusion)又称为信息融合(Information Fusion),旨在利用多种方式采集多源多模态的数据,得出更准确、统一的信息,辅助决策^[11]。①在国际上,以美国为代表的情报机构在“9·11”事件之前,烟囱式的情报体制阻碍了数据的融合,导致无法产生更高价值的情报产品。之后美国吸取教训,逐步建立多源融合中心,例如公开信息源中心,通过挖掘、融合和分析多源异构的公开数据,支持政府决策。②在国内,数据融合的研究多集中于多源数据理论、方法和技术的探讨。化柏林^[12]构建了科技大数据在情报服务中的应用框架,并指出多源异构的融合技术推动情报服务转向了主动服务、嵌入式服务。李广建等^[13]从知识融合的角度研究了多源异构数据的计算型情报分析。李纲等^[14]提出了科学社会计算的概念,解决了多源数据的组织与建模问题。但总体来看,有关多模态数据融合的研究多集中于计算机视觉、遥感等领域,在情报工作中的应用还有待深入。

综上,无论是用“海量数据”还是“大数据”来表征这个时代,情报感知过程中多源多模态数据的融合问

题都无法回避,然而现有研究大多聚焦于数据融合必要性、理论的探讨,融合的深度也浮于数据表面,并不能支撑细粒度的情报感知任务。破除数据异构、异质、多模态等特征给数据组织和关联带来的阻碍刻不容缓。

基于此,本文在系统梳理情报感知和多源多模态数据特点的基础上,阐述了多源异构数据的细粒度融合方法以及多模态数据的语义融合方法,并提出了时空数据辅助的多模态数据融合思路。旨在弥补信息不完备情况下的感知弊端,支撑情报感知由被动转向主动,促进情报感知的智慧发展。

2 基于多源多模态数据的情报感知

情报感知针对的是未知或知之不详的对象,情报工作者所面临的问题不确定性更强。因此,只有及时、准确地捕捉用户的情报需求,结合海量的数据去挑战更为潜在的假设,才能发现数据之间更为隐蔽的关联,完成好情报感知的任务。此外,过去的情报需求往往由情报用户直接提出,但随着情报工作的转型升级,如何主动感知潜在的情报用户需求也变得越来越重要。

在情报感知过程中,情报素材的处理也至关重要。本文所指的情报素材是多源、异构、多模态的数据,包括期刊、专利等科技文献数据,博客、论坛等新媒体数据,空间媒体、行为轨迹等空间数据。由于数据的更新迭代速度极快,使得情报感知从泛化的感知相关数据逐渐发展到更加细致化的感知分支数据^[15]。如何把这些多源多模态的数据融合起来,统一规范、消除冗余,形成数据的互补与交叉印证,提供统一的数据视图,值得我们关注。

综上,本文将重点梳理情报感知模型中,多源多模态的数据融合什么、怎么融合的问题,即多源多模态数据的细粒度融合与融合方法的研究(见图1)。



图1 情报感知模型

3 多源多模态数据的特征与挑战

3.1 多源多模态数据的特征

随着网络化、数字化的快速发展,互联网数据的规模不断扩大,来自不同来源的数据被收集在一个数据集中,这种类型的数据被称为多源数据;而多模态数据是指不同形态的数据形式,多指文本、图像、音视频等混合数据。两者之间存在着或强或弱的内容跨越和语义关联,一般来说多源的数据,也多涉及多模态的处理问题。

总体来说,互联网背景下的多源多模态数据主要有以下特征:①多源和多模态的数据混合并存,两者密不可分,在处理不同粒度的数据时,是否多源和多模态的判断都不是绝对的;②数据外部的组织结构复杂,有结构化、半结构化和非结构化等多种类型;③不同类型的数据从不同方面表达同一语义,例如有些数据虽然在底层上是异构的,但在顶层语义上却是一致的。

具体而言,在情报感知过程中所使用的多源多模态数据是指公开、可以合法获取的数据,特别是网络上的开源数据。在国际上,RAND公司将开源的数据划分为来自新闻媒体、灰色文献、社交媒体等的数据^[16];CIA则将其划分为来自互联网、新闻媒体、期刊、地理空间信息等类型的开源情报^[17]。因此,结合情报感知的特点,本文所指的多源多模态数据主要有:文本、图像、音频、视频、矢量、混合数据6种模态;从来源来看可以分为:科技文献数据(期刊、会议论

文、专利、基金报告等)、社交媒体数据、新闻报道数据和地理时空数据(经济人文数据、位置轨迹数据和空间媒体数据等)4种来源。

首先,文本数据的获取难度小、内容解析容易,但异构性强。例如以文本为主的科技文献数据中,期刊、专利等提供了摘要、关键词和全文等,而基金项目、行业报告等只提供了全文,如果从全文开始分析,当处理高维数据时,会因文本长度的过载使得效率降低^[18]。因此异构文本数据的融合粒度至关重要。其次,图像、音视频数据多来自社交媒体和新闻媒体,所蕴含的信息更加直观,但是如果将其应用于情报感知中,更重要的是深层语义的理解。只有将深层语义信息与感知任务结合,才能发挥多模态的互补优势。此外,由于其发布门槛低、操作简单,在使用时要注意对数据质量的评估。最后,矢量数据是指来自地理空间的数据,往往具有时间、空间和属性3方面的固有特征,呈现出时空紧耦合、数据高维、复杂语义关联的特点^[19]。

综上,公开的多源数据对于情报感知具有较强的适用性,但文本数据的应用必须解决其多源异构带来的分析阻碍,从不同粒度的融合出发,整合彼此增益的信息;图像、音视频数据的应用必须要理解深层的语义,将直观感觉转换为计算机可理解;矢量地理数据为情报感知的时序和地理分析提供了新的视角,可以根据感知任务的特点进行恰当运用(见图2)。

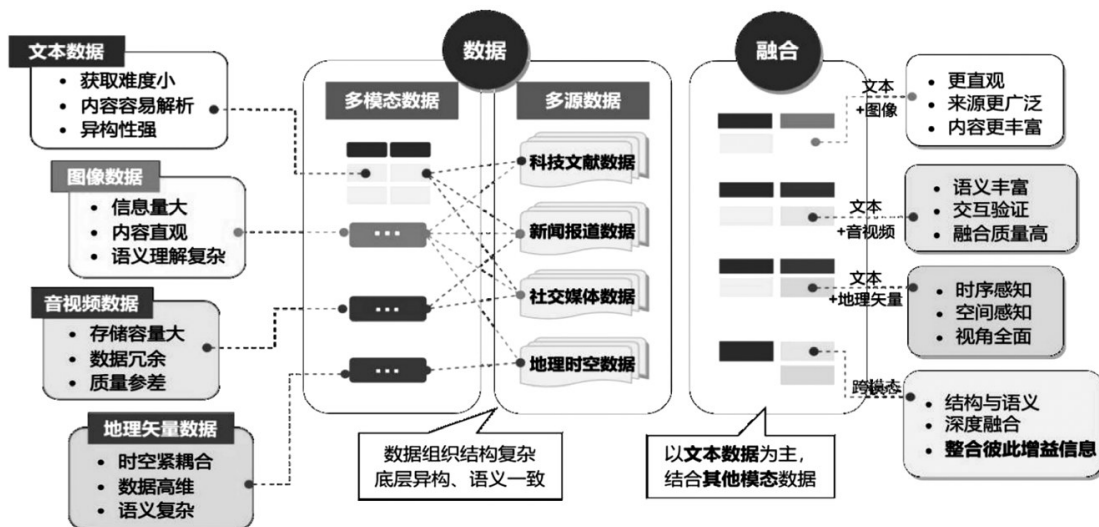


图2 多源多模态数据的特征

3.2 多源多模态数据面临的挑战

情报感知视角下情报工作所需要的数据支撑已经不再局限于传统的文献、事实型数据,网络开源数据所呈现出的上述特征对于提升数据的语义理解具有重要意义。如何打通并利用好这些多源多模态的数据,成为情报感知的重要基础。本文梳理了多源多模态数据在情报感知应用中亟待解决的挑战,旨在为后文数据的融合研究提供思想指导。

a. 数据不完备,数据所蕴含的碎片化信息使得情报线索呈现复杂性,实现数据的快速调动、文本挖掘与分析,发现潜在规律和联系至关重要;b. 数据质量良莠不齐,数据的质量直接影响了后续情报分析的成败,需要筛选数据自身质量高、情报价值高、对感知任务贡献度高的数据;c. 数据的动态调整,数据体量并不是越大越好,而是要随着情报感知任务的需求逐渐清晰和聚焦;d. 数据归属问题,要在数据安全、允许的法律规范下进行数据的获取、融合与分析。

4 面向情报感知的多源多模态数据融合方法研究

如何将互补但冗余的多源多模态数据有效融合,支撑情报感知,面临着以下挑战:①数据异构沟

沟和语义鸿沟亟待解决,不同源、不同类型的数据往往结构差异较大,且语义复杂;②数据融合粒度问题,如何根据不同的感知需求确定不同的融合粒度至关重要;③时序数据融合,海量的数据难免会引发快速查询等一系列问题,如何利用时间序列数据更快、更好地做出决策亟待解决;④智能融合,机器技术的蓬勃发展使“人+AI”的组合引人瞩目,需要在充分利用机器学习的基础上,更高效、智能的融合。

基于此,本文将从以下3个方面具体展开:①针对多源数据的异构鸿沟导致文本的相似性难以直接度量,融合粒度难以确定这一问题,聚焦单模态数据的融合,主要是指文本数据的多粒度融合。②针对多模态数据的语义鸿沟问题,重点调研了基于深度学习的多模态数据语义融合方法,特别是基于异质图神经网络的智能化融合方法。③情报感知中无论是时序数据还是空间数据,都能在情报感知的演化、预测类任务中提供全新的视角。因此应充分利用时空数据的交互,将时序、空间信息融入感知过程。

综上,面向情报感知的多源多模态数据融合模型(如图3),根据不同的情报感知任务特点,将数据的支持划分为单模态文本数据、多模态数据与地理时

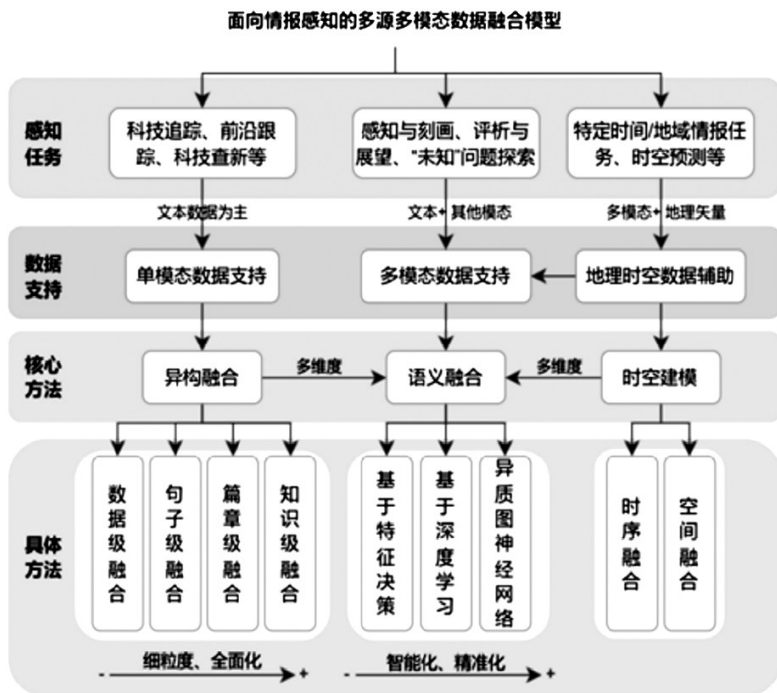


图3 面向情报感知的多源多模态数据融合模型

空辅助数据。具体的研究方法包括细粒度的数据级融合、句子级融合、篇章级融合和知识级融合,基于特征级、深度学习和异质图神经网络的语义融合,基于时空数据辅助的多源多模态数据融合。

4.1 多源异构数据的细粒度融合

目前学者们探讨的多源数据融合模型多从融合的过程出发,对各组成部分间的相互作用过程进行描述,例如Paul^[20]将融合过程划分为特征提取、位置融合、关系融合和关联融合4个步骤。还有学者从数据层、特征层和决策层融合3个层次概括多源数据融合模型。本文借鉴上述思路,结合情报感知任务的需求,认为面向情报感知的多源异构数据融合必须是细粒度的,在考虑融合过程和层次的基础上,是由简单的数据级组合融合扩展到考虑词粒度、短语粒度等的句子级融合,进而深化至考虑篇章语义、篇章结构、知识结构的细粒度融合,以揭示数据的深层规律,做出前瞻性强、价值高的判断。

4.1.1 数据级融合

数据作为情报感知的源头,可以帮助情报分析人员快速获取所需信息,其重要性不言而喻。多源异构数据的数据级融合可以分为数据直接融合和元数据融合两种不同的粒度。

数据级别直接融合。当情报感知的任务清晰,数据源结构相似、主题相似或核心数据明确时,可以将不同来源的数据在同一主题的指导下拼接组合,可以看作是数据集的扩大,是数据间的浅层融合^[21]。例如,数据A和数据B融合,补充数据C缺失的字段,或在数据A、B、C中选择高质量的字段。这种融合通常发生在学科前沿热点的识别、产业发展的变革规律、科研关系的发展演化等情报研究中,把同一主题、结构相似的数据直接融合,发现其演进的过程和规律,进而实现趋势的预测,早醒远眺。

元数据级别融合。元数据作为描述数据的数据,不仅能够 在情报感知中帮助情报人员快速获取所需素材,而且在对元数据进行抽取、映射、融合等操作时,能够有效避免数据的异构问题。主要有三种融合方式^[22]:①去重融合,当同一资源的元数据非空且一致时,只需要简单去重,保留一个来源的元数据。例如论文中的“关键词”元数据项均为“人工智

能”,则直接将其作为融合后的“题名”。②互补融合,当不同来源的元数据部分缺失时,需要对空值进行填补。例如某篇论文中的“作者”,各来源中只有一个来源的“作者”非空,则将非空项信息进行保留。③规则融合,如果某一资源各来源的元数据均非空,则保留符合元数据著录标准的数据。综上,去重融合和互补融合只针对元数据本身进行操作,简单快捷;而规则融合需要考虑元数据的质量和异构问题。

4.1.2 句子级融合

在情报感知过程中,情报素材的评判对感知的结果至关重要。例如,针对前沿动态跟踪类的感知项目,情报工作人员必须要全面、系统地获取可能与用户需求相匹配的数据,这时需要对含有“最新”“最前沿”等字样的文章极为在意,基于数据粒度的融合已经不能满足情报需求。而词语级和短语级融合大多以词作为语义单元,对于英文文档来说适用于这种融合粒度,但是中文的分词、语义等直接影响了融合结果,只考虑词语或短语级融合不一定能取得很好的效果。因此需要考虑到多语义的文档信息,综合考虑文本的词粒度信息、短语粒度信息和句粒度信息,通过更多层级和多粒度的融合,弥补数据级融合的信息损失,从而达到更好的融合效果。

在句子级融合的具体研究上,Luo等^[23]融合了句子级和文档级信息,提出了分层语义表示模型,并结合BiLSTM使模型充分利用上下文和全局信息。杨春霞等^[24]通过字、词、句的句法关系以及字-词关系,对字和词嵌入后进行多粒度融合,利用CNN提取语义特征及上下文特征完成了新闻文本的主题分类。由此可知,当数据级的融合不能支撑情报感知时,基于字、词、句的多粒度融合模型能更好的理解文本的实体信息和语义信息,同时兼顾文本的上下文信息,支撑情报感知任务。

4.1.3 篇章级融合

篇章级融合主要是从文章的逻辑结构和篇章单元之间的语义关系出发,挖掘更深层次的语义和结构信息。篇章级融合有助于理解篇章的中心思想和主旨内容,对情报感知时间线的构建、感知信息的抽取等任务具有重要意义。以情报感知中的科技政策

类任务为例,主题各异的政策看来在是离散的,在内容上缺乏相关性,而实际上却具有深层次的联系。在这时,就可以将不同主题文本中的关键信息抽取,以篇章级的事件时序关系或主题联系为特征,生成关于全局数据的新的更加抽象的特征,以此揭示数据的深层次规律。

在篇章级事件时序关系融合方面,事件的时序关系与篇章关系密切相关,如果仅以句子级的信息作为支撑,难以高效识别并融合事件的时序关系。戴倩雯等^[25]融合了句子的依存关系和篇章的修辞关系,利用BiLSTM捕获更多有效信息,较好地识别了事件的时序关系。刘耀等^[26]提出了融合篇章结构的文本知识网络,利用BiLSTM-CRF和SVM完成关系的标引,融合了篇章特征和句法语义特征。由此可知,融合过程中如果考虑了篇章级的时序关系,文章中的事件并非孤立的存在,使得全文的一致性得以保证,增强了情报感知的准确性。

篇章级主题融合的重点在于发现数据间的主题关联,找出背后的情报线索,以此洞察事务的本质。在具体方法上,基于LDA的主题模型在情报感知的演化类任务中具有较好的效果。徐硕等^[27]将主题、作者和时间关联提出作者主题演化模型(Author-Topic over Time, ATOT),挖掘了隐含主题和作者研究兴趣随时间变化的规律。冯佳等^[28]从载体-特征-关系3个层面构建了改进LDA的多源数据融合模型。但是LDA模型产生的主题之间相互孤立,会受到文本自身篇幅短、内容格式散乱等多方面的影响。随着深度学习的突破性应用,word2vec通过向量空间的相似度衡量文本的相似度,克服了词袋表示的不

足,使得语义信息更加丰富,则更适于从短文本中提取主题。

4.1.4 知识级融合

知识级融合中最具代表性的是应用于知识图谱的知识融合(Knowledge Fusion),即合并两个知识图谱中各层面的知识,将多源的有关同一实体的描述信息有效融合。相比于篇章级的融合,知识级的融合粒度更加细致,得到的结果也更加准确和全面。由于同一领域的数据具有各自的知识库,不同领域的数据进行交互时,知识的共享性、动态性决定了知识的构建过程中不可避免地会遇到知识异构的现象。南京大学计算机软件新技术国家重点实验室开发的众包知识融合系统FactChain^[29],解决了知识因时效性更新、多源知识间冲突的问题,为知识的多源共享与融合提供了指导。因此,解决知识级的融合问题,能更好地支持细粒度的情报感知,将知识异构现象冰解冻释。

在知识融合中常用的工具主要包括Falcon-AO、Dedupe、Limes和Silk等。其中,Falcon-AO^[30]是一个本体自动匹配系统,是目前web本体匹配最为实用和流行的方法之一。而Dedupe^[31]可以支持多种数据类型和自定义类型,为多源异构数据的融合提供了合适的选择。Limes是基于度量空间的框架,适用于大规模的数据链接;Silk是集成异构数据的开源框架,提供了图形化的用户界面,可以更加方便的进行记录链接。

综上,异构数据融合,特别是异构文本数据融合时,基于数据级、句子级、篇章级和知识级的融合各有利弊(见图4),但在情报感知的应用时,要根据情报

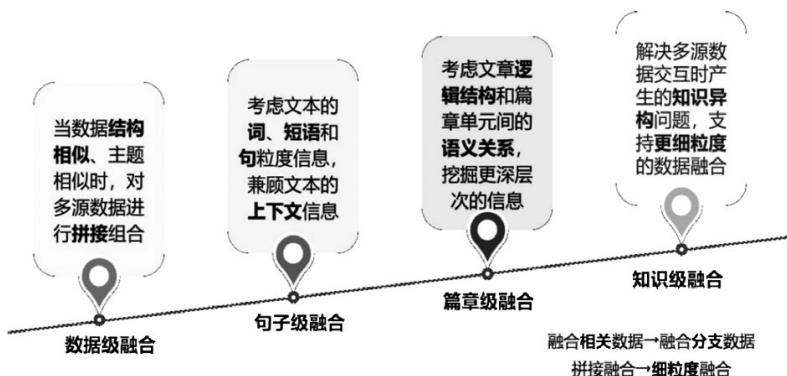


图4 多源异构数据融合的发展趋势

任务的特点,选择恰当的融合粒度,并在必要的时候对不同粒度的融合进行组合,促进情报感知的智能化和精准化。

4.2 多模态数据的语义融合方法研究

随着智能设备的更新迭代,越来越多的人在微博、微信、抖音等社交平台上以文本、图片、音视频等形式发表个人观点。面对海量的社交媒体和新闻报道数据,若能实现其合理利用,将在情报感知中产生巨大的价值。相比于单模态的文本数据融合,多模态数据融合则更加重视对数据深层次语义的理解,以及不同数据间关系的建模,有时也通过抽取不同模态的交互信息,高效利用海量的媒体数据。

总体。Omnivore采用统一的预训练模型处理图片、视频和3D图像三种视觉模态的输入,并映射到统一的特征空间;Data2vec则采用了统一的模型和训练框架,分别去训练文本、图像和音频,是第一个适用于文本、语音和图像等多领域的高性能自监督算法。此外,Bridge-Tower(MSRA)也提出了多模态预训练模型新的特征融合方式,其在单模态表示学习的中间层进行跨模态的交互,不仅可以有效应对模态缺失的问题,而且能更好的学习多模态的交互。

4.2.1 基于特征、决策级的多模态数据融合

关于多模态数据融合的智能处理,有特征级和决策级2个层次的融合方法。特征级融合适用于不同模态的数据高度相关的情况,采用提取特征信息融合的方式^[33]。Yang等^[34]采用了特征级的多模态数据融合,将音频和视频数据应用于语音识别,但特征提取难度大,效果有待提高。此外,最初应用于文本分类问题的PLSA(Probabilistic Latent Semantic Analysis, PLSA)模型近年来被应用于图像、音视频等多媒体数据的分析,并取得了不错的效果。例如,熊回香等^[35]利用PLSA模型和潜在主题变量的不同特征,构建了潜在主题语义空间,进而发现多模态数据间的相关性,克服了文字潜在语义信息的问题。在决策级融合方面,谢豪等^[36]构建了多层语义决策级融合模型,充分挖掘图文两种模态数据之间的相似性信息,对社交媒体的图文进行了情感分类。但需要注意的是,决策级的融合在处理异质的多模态数据时,会产生信息的损失。

4.2.2 基于深度学习的跨模态数据融合

随着人工智能的发展,跨模态的数据融合逐渐走向人们的视野。跨模态数据融合通过关系抽取来获得不同模态间的交互信息,再结合数据自身得到相互独立的特征,比多模态数据融合所得到的特征更加全面和精确。

a. 基于CNN的跨模态数据融合。卷积神经网络(Convolutional Neural Network, CNN)不需要手动提取数据的特征,可以共享权值并直接输入网络,被广泛应用于图像、视频等模态数据的处理中。2015年You等^[37]首次提出拼接图文特征以实现跨模态分类,使用ImageNet预训练的CNN提取图像特征,Word2Vec提取文本特征,此后的研究均是该方法的改进。申自强^[38]利用CNN和CBOW(Continuous Bag of Words, CBOW)分别提取图文特征,但也只能提取高级语义特征。针对这一问题,陈巧红等^[39]利用VGG13网络外接卷积层,提取3个层次的图像特征,对文本数据和图像数据进行跨模态融合。Hu等^[40]利用预训练的CNN提取图像特征,选择LSTM在Glove词向量上提取文本特征表示。

b. 基于注意力的跨模态数据融合。在深度学习中引入注意力机制(Attention Mechanism)后,使得机器更贴合人类行为,也在多模态数据融合中取得了一定效果。Xu等^[41]基于注意力机制分别计算了图文之间的双向注意力,将文本实体和图像语义对齐,实现了分层次的跨模态融合。陈小敏等^[42]提出了基于Attention的TAI-CNN跨模态情感分类方法。Yu等^[43]利用多头注意力机制(Multi-Head Attention Mechanism)将图像特征融入文本特征中,并利用自注意力机制提取文本数据的上下文信息,使跨模态的融合更加清晰。在此基础上,又提出视觉注意力网络(Visual Aspect Attention Network)以解决图像模态的数据为主的应用场景,更精确的进行多模态数据融合。综上,跨模态数据融合模型不再局限于简单的特征提取、决策融合或单一神经网络,而是可以根据实际问题的需要和应用环境,选择适合于当前情报感知任务或感知对象的模型。

4.2.3 基于异质图神经网络的多模态数据融合

情报感知中的多模态数据大都具有图结构,图

神经网络 GNN(Graph Neural Network, GNN)^[44]可以通过图节点传递和捕捉数据的相关性,推动了数据融合领域相关研究的发展。然而传统的 GNN 模型,只能处理某一类型节点构成的规则拓扑图,也就是说只能表示同一模态或类型的数据,但是在实际的情报感知过程中,多模态的数据并非只有一种类型的节点构成,传统 GNN 的模型质量明显下降。

近年来,由不同类型的节点和边构成的异质信息网络(Heterogeneous Information Network, HIN)出现,弥补了同构网络的不足^[45]。越来越多的学者将多模态的交互信息建模为异质信息网络,保留更加全面的语义信息,完成了更精确的知识关联与发现。特别是面对情报感知中的海量、多源、异构、多模态的数据,异质信息网络的融合优势使其成为解决多模态数据融合的重要途径。

既然异质信息网络可以融合丰富的结构和语义信息,解决数据异质融合的挑战,那文字、图像、时空数据等模态数据是否可以采用异质信息网络和图神经网络结合的方式进行建模与分析?陈卓等^[46]将 GNN 与异质信息网络结合,对多源异质数据进行了解码,实现了用户项目和社交数据两种网络节点的有效融合。冯小东等^[47]基于异构图神经网络实现了社交媒体用户与信息的交互建模与融合。Hu 等^[48]利用异质信息网络、GNN 和 LSTM 完成了用户、新闻、新闻主题间的模态交互,捕捉了用户的兴趣,对于情报感知的用户追踪类任务具有一定的启示意义。在以图像模态数据为主的领域,基于异质图神经网络

的建模方式也有初步尝试。Li 等^[49]将图像中的位置、空间、语义等信息建模为异质网络,并利用图注意力机制完善了全局建模。因此针对情报感知中的多模态数据,基于异质图神经网络的数据融合是具有一定的竞争优势的(见图 5)。

4.3 时空数据辅助的多源多模态数据融合

在情报感知时,允许不同模态的数据之间互融互通、彼此增益是多源多模态数据融合面临的挑战,如果仅依靠深度学习等数据挖掘的方法,情报用户难以对数据和感知的过程本身进行直观的理解和探索。基于时间数据和空间数据辅助的多源多模态数据的融合,可以充分利用时空的交互,将时序数据、地理空间信息融入整个数据分析、感知决策的过程^[50]。

时序数据融合。时序数据表示着事件发生的先后顺序,串联了情报感知中各个事件或要素的发展演化,一些看似离散着被叙述的事件,往往背后存在着一定的联系。如果能及时、准确地抽取并融合情报感知数据源中事件的时序关系,对于情报感知的脉络梳理起到了很好的推动作用。因此,时序数据的融合也被看作是情报感知中的一项重要任务,受到了越来越多的关注。

空间数据融合。随着全球定位、无线传感器等技术的发展,产生了海量 GPS 数据,进而催发了大量基于位置服务(Location-Based Services, LBS)的应用平台。典型的 LBS 的社交网络,如 Twitter、新浪、高德等运营过程中,产生了海量的时空语义数据^[51],助力了情报的多维度感知。Twitter 每天约产生 1000 万

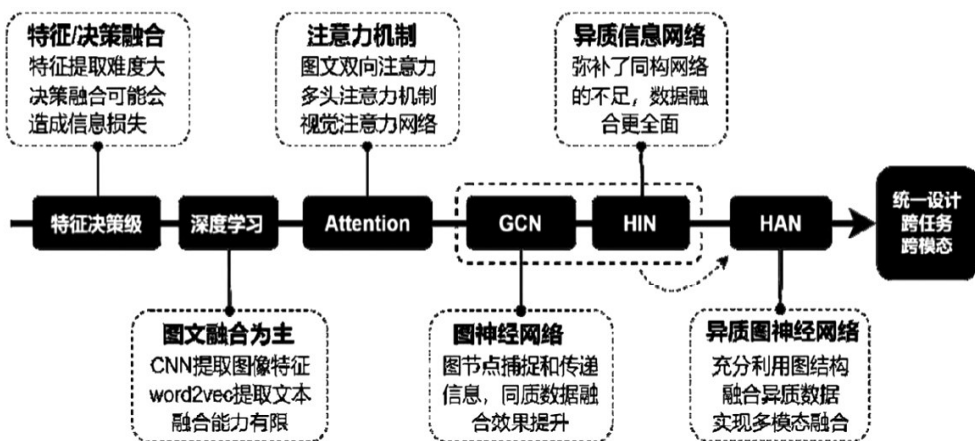


图 5 多模态数据融合的发展趋势

条带有地理信息的数据;美国 Foursquare 中更是存储了超 100 亿条的位置签到信息^[52]。在情报感知中,如果将这些数据的 GPS 信息与多模态数据结合,既可以表示感知对象在一定时间范围内的位置变化,也可以从细粒度、多层面和多视角记录感知对象的变化状态。

5 总结与展望

本文在系统梳理情报感知与多源多模态数据特点的基础上,提出了多源多模态数据在情报感知应用中存在的问题,并在此基础上详细梳理了细粒度的数据融合方法。针对单模态数据特别是多源异构的文本数据,根据情报任务的不同选择数据级、句子级、篇章级和知识级的融合方法。面对复杂的多模态数据,基于深度学习的方法始终是语义融合的首选,值得一提的是基于异质图神经网络的融合方法,是未来研究的重点。最后,提出了时空数据辅助的多模态数据融合思路,为情报感知提供了新的视角,未来一段时间,融合时序数据或全时空大数据的情报感知研究仍具有重要的研究意义。

总体而言,情报感知视角下,多粒度的多源、异构、多模态数据融合不可避免。在实际应用时也需要注意不同融合粒度之间的选择与组合,以及各模态数据的综合、准确应用。特别是面对异质、交互的多模态数据,基于异质图神经网络的融合方法实力强劲,但也存在不少难题,例如从多模态数据中提取恰当的对象和关系;实际应用中的异质网络往往动态变化、规模巨大、模式丰富;异质图网络中,局部和全局信息的关系如何把握等都亟待解决。

参考文献:

- [1]王延飞,何芳,闫志开.情报研究方法构建的关系基础[J].情报杂志,2015,34(4):1-3,26.
- [2]王延飞.杜元清.情报感知论[M].北京:科学技术文献出版社,2021:2-3.
- [3]Liu J, Li T, Xie P, et al. Urban big data fusion based on deep learning: An overview[J]. Information Fusion, 2020, 53: 123-133.
- [4]裘惠麟,邵波.多源数据环境下科研热点识别方法研究[J].图书情报工作,2020,64(5):78-88.

- [5]杜元清.科技情报工作的特点和规律[J].中国国防科技信息学会通讯,2008(2):32-56.
- [6]杨峰,姚乐野,范炜.情景嵌入的突发事件情报感知:资源基础与实现路径[J].情报资料工作,2016(2):39-44.
- [7]王延飞,赵柯然,陈美华,等.情报感知的研究解析[J].情报理论与实践,2018,41(8):1-4.
- [8]杨峰,张月琴,姚乐野.基于情景相似度的突发事件情报感知实现方法[J].情报学报,2019,38(5):525-533.
- [9]马红岩,陈峰,曾文.科技情报中多源信息融合的模式构建[J].中国科技资源导刊,2022,54(3):1-8.
- [10]白如江,张玉洁,赵梦梦,等.面向关联推理的智慧情报感知:内涵、组织与路径[J].情报理论与实践,2022,45(8):31-37,67.
- [11]李广建,化柏林.大数据分析 with 情报分析关系辨析[J].中国图书馆学报,2014,40(5):14-22.
- [12]化柏林.科技信息大数据在情报研究服务中的应用[J].图书情报工作,2017,61(16):150-156.
- [13]李广建,罗立群.走向知识融合——大数据环境下情报学的发展趋势[J].中国图书馆学报,2020,46(6):26-40.
- [14]马亚雪,毛进,李纲.面向科学社会计算的数据组织与建模方法[J].中国图书馆学报,2021,47(1):76-87.
- [15]周京艳,刘如,张惠娜,等.基于观点的情报感知方法研究[J].情报杂志,2019,38(12):35-39,47.
- [16]Williams H J, Blum I. Defining second generation open source intelligence (OSINT) for the defense enterprise[EB/OL]. [2022- 12- 15]. https://www.rand.org/pubs/research_reports/RR1964.html.
- [17]Central Intelligence Agency. Intelligence: Open source intelligence[EB/OL]. [2022- 12- 15]. <https://www.cia.gov/news-information/featured-story-archive/2010-featured-story-archive/open-source-intelligence.html>.
- [18]张维冲,王芳,赵洪.多源信息融合用于新兴技术发展趋势识别——以区块链为例[J].情报学报,2019,38(11):1166-1176.
- [19]李德仁.展望大数据时代的地球空间信息学[J].测绘学报,2016,45(4):379-384.
- [20]Paul L F. Sensor datafusion[J]. Journal of Intelligent and Robotic Systems, 1988, 1(2): 103-116.
- [21]李志勇,孟潇,杨海丽,等.基于信息融合模型的科技情报质量控制研究[J].情报杂志,2019,38(1):54-60.
- [22]李静,胡潜,李想,等.基于决策树的多源文献元数据融合研究[J].图书情报工作,2022,66(6):118-125.
- [23]Luo Y, Xiao F S, Zhao H. Hierarchical contextualized

representation for named entity recognition[C]//Proceedings of the 2020 AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2020: 8441-8448.

[24]杨春霞,李锐,秦家鹏.一种粒度融合的新闻文本主题分类模型[J].小型微型计算机系统,2020,41(11):2256-2259.

[25]戴倩雯,张龙印,孔芳.融合依存关系和篇章修辞关系的事件时序关系识别[J].模式识别与人工智能,2019,32(12):1100-1106.

[26]刘耀,张越,叶璐.融合篇章结构的文本知识网络构建[J].图书情报工作,2021,65(21):118-130.

[27]Xu S, Shi Q, Qiao X, et al. Author-topic over time (AToT): A dynamic users' interest model[J]. Lecture notes in electrical engineering, 2014, 274: 239-245.

[28]冯佳,穆晓敏,王伟.面向研究前沿识别的载体-特征-关系融合模型研究[J].图书馆杂志,2020,39(9):56-63.

[29]朱向荣,吴鸿祐,胡伟.FactChain:一个基于区块链的众包知识融合系统[J].软件学报,2022,33(10):3546-3564.

[30]知识图谱生命周期及工具[EB/OL].[2022-10-30]. <https://blog.csdn.net/getinglxf/article/details/103955259>.

[31]OpenKG.Dedupe: 知识链接 python 库[EB/OL].[2022-10-31].<http://www.openkg.cn/tool/dedupe>.

[32]清华大学周伯文:从原则到实践解读多模态人工智能进展与可信AI[EB/OL].[2022-10-06].<https://baijiahao.baidu.com/s?id=1744383892849739250&wfr=spider&for=pc>.

[33]祁铎颖,贺萍.跨模态数据融合综述[J].软件工程,2022,25(10):1-7.

[34]Yang X C, Feng S, Wang D L, et al. Image-text multi-modal emotion classification via multi-view attentional network[J]. IEEE Transactions on Multimedia, 2021, 23: 4014-4026.

[35]熊回香,杨滋荣,蒋武轩.跨媒体知识图谱构建中多模态数据语义相关性研究[J].情报理论与实践,2019,42(2):13-18,24.

[36]谢豪,毛进,李纲.基于多层语义融合的图文信息情感分类研究[J].数据分析与知识发现,2021,5(6):103-114.

[37]You Q, Luo J, Jin H, et al. Joint Visual-Textual sentiment analysis with deep neural networks[C]//Proceedings of the 23rd Annual ACM Conference on Multimedia Conference. ACM, 2015: 1071-1074.

[38]申自强.一种基于图文融合的跨模态社交媒体情感分

析方法[J].软件导刊,2019,18(1):9-13,16.

[39]陈巧红,孙佳锦,孙麒,等.基于多层跨模态注意力融合的图文情感分析[J].浙江理工大学学报(自然科学版),2021,47(1):85-94.

[40]Hu A, Flaxman S R. Multimodal sentiment analysis to explore the structure of emotions[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2018: 350-358.

[41]Xu J, Huang F, Zhang X, et al. Visual-Textual sentiment classification with Bi-Directional Multi-Level Attention Networks[J]. Knowledge-Based Systems, 2019, 178: 61-73.

[42]陈小敏,许华虎,方迪恺,等.基于注意力机制的TAI-CNN图文跨模态情感分类方法[J].计算机应用与软件,2021,38(5):196-200,293.

[43]Yu J, Jiang J. Adapting BERT for Target-Oriented multi-modal sentiment classification[C]//Proceedings of the 28th International Joint Conference on Artificial Intelligence. IJCAI, 2019: 5408-5414.

[44]Wu Z H, Pan S R, Chen F W, et al. A comprehensive survey on graph neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(1): 4-24.

[45]石川,王睿嘉,王啸.异质信息网络分析与应用综述[J].软件学报,2022,33(2):598-621.

[46]陈卓,李涵,杜军威.基于异质图神经网络的推荐算法研究[J].湖南大学学报(自然科学版),2021,48(10):137-144.

[47]冯小东,惠康欣.基于异构图神经网络的社交媒体文本主题聚类[J].数据分析与知识发现,2022,6(10):9-19.

[48]Hu L, Li C, Shi C, et al. Graph neural news recommendation with long-term and short-term interest modeling[J]. Information Processing & Management, 2020, 57(2): 102142.

[49]Li L, Gan Z, Cheng Y, et al. Relation-aware graph attention network for visual question answering[EB/OL].[2022-10-12]. <https://arxiv.org/abs/1903.12314>.

[50]朱庆,付萧.多模态时空大数据可视分析方法综述[J].测绘学报,2017,46(10):1672-1677.

[51]朱少杰.多模态时空语义轨迹预测[D].徐州:中国矿业大学,2021.

[52]姚迪,张超,黄建辉,等.时空数据语义理解:技术与应用[J].软件学报,2018,29(7):2018-2045.