客户RL信息表-（研究生）

|  |  |  |  |
| --- | --- | --- | --- |
| 亲爱的同学：  为了更好地了解你，完成一套最适合你的文书申请材料，我们特制作了此套信息搜集表。此套信息搜集表一共有三个部分，即《客户信息表》，《客户SP信息表》，《客户RL信息表》。  本表属于三个部分中的《客户RL信息表》，表格中的信息将用在你的留学文书RL中，推荐信的作用是从第三者角度对申请者CV和SP及成绩的确认和重新解释，对申请者的陈述和成绩进行补充，老师的推荐是以同行的身份向录取委员会的教授提供自己的看法(尤其是在校教授的推荐信)，包括申请者的素质，动机以及申请者是否能从该课程获益等。  所以，作为录取委员会对申请者的考察标准之一，推荐信的作用是不可忽视的，好的推荐信会帮助您拿到学校的取录和奖学金，而差的，尤其是那些泛泛而论的，内容混乱的推荐信往往会导致申请功亏一篑，甚至会造成致命伤害。  除此之外，作为理工科的学生，虽然留学申请中GPA、语言等硬件条件很重要，但是想要得到教授的青睐，个人的学术能力、研究或实验项目也是不可小觑的一方面，所以请你一定要认真填写下面的每一项，把你的能力在实验项目中表现出来——让对方相信你是有能力，值得培养，能带来新元素的优秀学生。  我留学，我参与，我成长  让我们一同协作，共同完成一套最适合你的文书申请材料，筑起腾飞的梦想！  信息搜集表-填写指南   1. 因为本信息问卷是适用于理工研究生申请的，所以体现申请者的研究能力和创新能力尤其重要，所以在提供素材中一定要认真回顾和推荐人之间的交互，把自己的能力体现出来。 2. 你提供的素材越具体越详实，对你的申请越有利。 3. 请务必详实填写以下推荐信写作信息搜集表，避免过度笼统和陈词滥调，插入具体事例，避免堆砌泛泛的赞美之词。 4. 申请硕士课程的学生选择的推荐人一般为专业课老师，系主任/学院院长， 辅导员等，学生可根据自身情况选择三位推荐人。如跨专业申请，请选择与申请专业相关的老师。 5. 推荐人身份最好区别开，比如系主任，校领导，授课老师，实验室老师，工作或实习单位领导等等，要从不同的角度来对申请者的优点，能力进行评价。另外申请者要注意在填写推荐原因的时候要注意写出一到两个实例，不要写大话空话，这些实例要突出申请者的能力，要贴近专业。而且在提供信息时要站在推荐人的角度来思考。 | | | |
| **\*若知道专有名词的官方英文，请提供英文** | | | |
| **推荐人信息** | | | |
| **姓名：** | 王伟平 | **工作邮箱：(edu.)** |  |
| **职务：（例院长）** |  | **工作电话：** |  |
| **职称:（例教授）** | 教授 | **与你的关系：** | 暑研项目指导老师 |
| **工作单位：** | 中南大学计算机学院 | **工作地址：** |  |
| 1. **以时间为顺序，简述推荐人和申请者的关系发展，若有多种认识途径，请一一列出，并说明你们的熟悉程度。** | | | |
| 2023年暑假我联系王老师，申请做她的暑研。我被安排到“超大规模网络中恶意流量只能处置技术”课题组，负责网络威胁情报（CTI）的知识库的构建和基于CTI报告构建知识图谱的代码工作，以突破细粒度攻防知识图谱构建，为探索基于知识辅助决策的恶意流量智能处置新技术  打下基础 | | | |
| 1. **推荐人对本人哪方面能力留下深刻印象并且非常认可？因何事件体现出来？（请至少选2个例子）** | | | |
| 这个项目我主要负责知识图谱构建的代码实现。也就是说，要从海量的CTI报告中，抽取出其中提到的不同类型（type）的STIX（Structured Threat Information eXpression）实体（entity），并提取出一篇报告中不同实体之间的关系，并以此构建知识图谱。其中的难点在于：  1.不同报告的写作风格自然大相径庭，如何使模型能准确的适应新报告的写作风格，提取出其中的实体并进行关系分析。  2. 由于现有的实体知识库（knowledge base）有限，该如何解决面对一篇新报告，提取出没有包括在知识库中的新实体。  3.目前对长距离实体间关系的提取还没有非常有效的办法（如报告开头的实体与结尾出现的实体之间的关系）。  这个工作本质上可以视为一个NLP任务，而这也恰恰是我比较熟悉。首先，我快速地学习了有关CTI和STIX的相关知识，阅读了近期比较具有代表性的相关领域的文章，对此次任务有了一个清晰的认识。针对上述难点，我分别进行了如下应对措施：  1.首先将不同类型的实体进行分类：对那些有着清晰结构的实体，如Indicators Of Compromise（IOC），采用基于规则的算法（rule-based）进行提取；而对于没有清晰结构，但已经被归纳进知识库的实体，采用AC自动机算法（Aho-Corasick algorithm）进行多模式字符串匹配（multi-pattern string matching）；同时为了减少误报（false positives），结合了词性标注（Part-Of-Speech Tagging）的NLP处理技术。这样便能将找到的实体与知识库中的实体匹配成功。  2. 通过分析不同供应商和来源的报告，最后便可以确定引入新实体的有限方式，从而编写NLP处理可以识别的模式规则。结合上面提到的词性标注结果，采用Spacy（a free, open-source library for advanced NLP in Python）构建报告中每句话的依赖关系图（dependency graph），基于识别出来的模式规则进行新实体提取。  3.采用Sentence BERT模型，计算对知识库中的STIX关系对象（STIX Relationship Object，SRO）和从报告中提取的句子的嵌入(embeddings),然后，计算这些嵌入之间的余弦相似度，并对其进行归一化，以将其用作置信度值（confidence value）。这样，置信度超过阈值（threshold）就可以确定实体间的关系了。  此次任务过程基本全部由我独立完成，充分体现出我的代码能力和创新能力。我的模型也将被用于该课题以后的发展中，成为了该课题的基石，为恶意流量智能处置新技术提供知识辅助决策的必须材料（知识图谱）。  同时，此次任务也遇到许多困难。如在实现难点2的最后，我发现，光凭基于规则的算法进行新实体提取还是不够的：报告的写作风格千奇百怪，因此提取出所有写作风格对应的风格是不现实的，只能做到尽量囊括大部分；而这样无疑会引入了漏判。因此，我和组内另一位同级学生一起合作，我实现基于规则的算法，他实现基于机器学习、深度学习的算法，最后将我们的成果进行结合，先使用基于规则算法进行筛选，再通过机器学习算法查漏补缺，这样既能解决漏判问题，又能解决机器学习算法资源需求大的问题。充分体现了我的合作能力。  角度：**科研能力**（未知领域的探索, 理论应用实践, 发现分析解决问题，等等）**，创新能力，语言表达能力**（口语，写作）**，团队协作能力，领导才能**（决策能力，分析判断能力，激励能力，指挥能力）**，智力潜能**（独立思考能力，领悟能力，快速学习能力，等等） | | | |
| 1. **补充信息：** | | | |
|  | | | |
| 1. **推荐人背景介绍（留学或交流经历，国际上注明专业协会会员重点写）** | | | |
| 王伟平，中南大学计算机学院，教授，博士生导师，教育部网络空间安全教学指导委员会委员。中南大学计算机学院“网络威胁发现与隐私数据保护”学科方向负责人，湖南省网络空间安全研究生创新培养基地负责人。1991年获东南大学电子工程学士学位，1994年获中南大学计算机应用技术硕士学位，2004年获中南大学计算机应用技术博士学位。1994年开始在中南大学任教，2009年至2010年在美国德州A&M大学访学一年。2009年入选教育部新世纪优秀人才支持计划。主要研究兴趣包括互联网应用安全、物联网安全、安全大数据分析、网络协议分析与行为监测等。作为项目负责人承担了4项国家自然科学基金，承担及参与了其他重大科研项目10余项。在包括《IEEE Transactions on Information Forensics & Security》、《软件学报》、《计算机学报》等在内的国内外刊物发表学术论文90余篇。获得发明专利5项，国家软件著作权3项。作为主要参与者获得教育部自然科学一等奖1项，湖南省科技进步二等奖2项。  长期致力于信息安全人才培养，信息安全专业负责人，主讲课程：《网络安全》、《网络与系统安全》、《信息安全专业导论》、《计算机网络》等。担任教育部网络空间安全教学指导委员会委员、湖南省网络空间安全协会专家咨询委员会专家、中国网络空间安全人才教育联盟中南大学代表。  主页：[中南大学 王伟平 (csu.edu.cn)](https://faculty.csu.edu.cn/wangweiping/zh_CN/index.htm) | | | |