引言

随着互联网技术的发展与web 2.0时代的到来，Ajax（Asynchronous JavaScript And XML）的web应用发展，XML（Extensible Markup Language）作为传统的web数据交换与存储格式，因其语法复杂、数据冗余大和不易解析等问题，影响着网页的响应速度与数据传输效率；人们开始寻找另外一种合适的数据格式，因此JSON（JavaScript Object Notation）作为一种轻量级的数据交换格式，其具有与XML相同的可读性、可扩张性等方面的优点，也具有多兼容性、解析容易、结构简单等独特的优点，JSON引入Ajax通信之中成为研究热点；【屈展,李婵.JSON在Ajax数据交换中的应用研究[J].西安石油大学学报(自然科学版),2011,26(01):95-98+122.】同时，JSON与XML一样支持对异构数据的存储，为信息集成提供了新的解决方案，对结构化数据、半结构化数据以及无结构化数据均可转换成JSON结构；【倪睿熙.一种基于JSON的异构数据查询方法[J].无线电通信技术,2013,39(01):73-76.】JSON大量应用于Web信息表示与数据交换，数据传输效率与解析难度远优于XML，并在未来超越XML，Ajax也会变成Ajaj（Asynchronous JavaScript And JSON），成为流行的数据交换格式与数据存储格式应用于Ajaj中。【张沪寅,屈乾松,胡瑞芸.基于JSON的数据交换模型[J].计算机工程与设计,2015,36(12):3380-3384.】

Web已经是人们获取信息的重要手段，随着 JSON的加入，大量异构数据集成在JSON文档之中；而Web的普通用户对这些JSON文档的结构与内容并不了解，或者用户表述的查询意图是模糊或不明确的，因此提交的查询条件大多是用户自身模糊意图的模糊描述，称之为JSON模糊查询。例如，对于示例的二手房数据库（价格、面积、房龄、城市、卧室、卫生间、阳台、楼层…），用户可因自身的不确定而输入模糊查询条件：城市为南京，价格比较低，楼层在3层~6层之间，面积在100平方米左右，卧室至少3间，房龄最多3年；该模糊查询条件包含了模糊关系“最多”、“至少”等以及模糊词“低”。而在现有的JSON查询方式中，如JSONSQL、JSONPath等，仅能支持精确查询匹配，还没有能直接支持对上述的模糊关系与模糊词的查询；因此需要对上述的模糊查询条件转换为精确查询条件，而后通过JSON的精确查询方式以获取模糊查询的结果。

近些年来，在数据查询研究工作中越来越注重对精确查询条件的柔性化处理，基于关系数据库的不精确查询方法中，以相似度为基础，对查询标准进行扩展，查询标准的扩展程度由查询条件的重要程度决定【Nambiar U, Kambhampati S. Answering imprecise queries over autonomous web databases.[C] //Proceedings of the 22thInternational Conference on Data Engineering . Atlanta: IEEE Computer Society，2006：45-54】；但是，这些研究没有考虑用户直接提出的模糊查询条件，无法处理用户表达的模糊查询要求；对于模糊数据的处理，在由Zadeh提出模糊集【ZADEH L A, Fuzzy Sets. Information and Control . 1965】之后，模糊集就是用来作为模糊数据处理的理论依据和方法。在数据库领域，基于模糊集处理关系数据库中的模糊查询条件最早由Tahani提出，【Tahani V. A conceptual framework for fuzzy query processing: a step toward very intelligent database systems[J]. Information Processing Management ，1977,13(2):289-303】而后在模糊查询条件转换规则的基础上，Ma等人【孟祥福,马宗民,严丽,张霄雁.一种基于知识的关系数据库模糊查询方法[J].东北大学学报(自然科学版),2009,30(03):349-353.】提出结合领域知识的支持，增加用户偏好，实现在关系数据库之中的模糊查询。在XML数据查询中，利用本体处理XML文档的柔性查询，然后通过使用查询重写机制获取查询的近似结果，该过程需要领域专家参与建立本体知识库以及映射规则；而对于用户的模糊查询，文献【孟祥福,张霄雁,马宗民,彭晏飞.一种基于领域知识的XML数据模糊查询[J].智能系统学报,2012,7(06):525-535.】提出使用领域知识与模糊集结合实现对XML数据的查询条件转换，增加查询条件中查询谓词的重要程度和用户偏好对模糊查询扩展程度，并通过满足程度对查询结果进行排序，实现XML数据的模糊查询；文献【闫威,马宗民,严丽,王星.基于XML内容和结构的模糊查询[J].东北大学学报(自然科学版),2011,32(07):931-934.】通过对XPath查询表达式的模糊扩展，以模糊集为基础，采用模糊查询松弛与查询结果排序打分方法实现对XML数据进行模糊查询。在JSON数据查询领域，目前研究的热点在于基于JSON结构特点对异构数据的查询，对于用户直接表达的模糊查询要求未能进行处理，也没有对JSON模糊查询的转换处理。因此，本系统基于模糊集理论，结合领域知识库设计JSON数据的模糊查询扩展，实现对用户直接输入的模糊查询要求进行处理；此系统可对用户输入的模糊关系和模糊词以及对精确数据区间进行JSON数据查询的模糊扩展，不改变原有的JSON查询方式，并对所有JSON文档包括异构数据结构具有通用性，为JSON模糊查询实现有效的查询扩展方式。

JSON

JSON（JavaScript Object Notation）是一种轻量级的数据交换格式，采用完全独立于语言的文本格式，基于JavaScript语言的子集，使用一种具有嵌套数据元素的结构，易于人们的阅读与编写，同时也便于机器的解析和生成。【Introducing JSON[EB/OL].http://json.org】JSON构建有两种结构：“名称/值”对的集合（A collection of name/value pairs）与值的有序列表（An ordered list of values），在不同语言中，前者可以理解为对象（object）、记录（record）、结构（struct）、字典（dictionary）、哈希表（hash table）、有键列表（keyed list），后者可理解为数组（array）；因此在同样基于这些结构的编程语言之间，可使用JSON作为他们的数据交换的格式。

JSON具有多种形式，其中对象（object）和数组（array）是最常见的两种形式；前者对象是无序的“‘名称/值’”对集合，以“{”开始，以“}”结束，每个属性名称和值之间使用“:”,属性名称之间使用“,”分隔；



图1 JSON对象结构形式

后者数组是值的有序集合，以“[”开始，以“]”结束，每一个值之间使用“,”分隔。



图2 JSON数组结构形式

结合对象形式和数组形式的简单JSON数据结构：

{

“FirstName”:”John”,

“LastName”:”Smith”,

“PhoneNumber”:[

{

“type”:”home”,

“number”: "212 555-1234"

},

{

“type”:”office”,

“number”: "646 555-4567"

}

]

}

模糊集

模糊集的定义