Research on the Construction Technology of Knowledge Graph in Aviation

Xue Wang*, Xizhong Yang, Jiapan Fu and Xuyi Qiu

Science and Technology on Avionics Integration Laboratory, 432 Guiping Road, Xuhui District, Shanghai, China

*E-mail: snow_wx@163.com

Abstract. In order to alleviate the overload of airspace information under the network environment and integrate the scattered and useful airspace information, the technical architecture of knowledge graph in the aviation field is proposed. This paper describes the technical architecture of aviation knowledge graph construction, including knowledge extraction, knowledge fusion, knowledge processing, knowledge update and other key technologies and applications. Finally, an instance of knowledge graph for air combat is constructed. The knowledge graph discovers the hidden relationship between task and maneuver through resource sharing and information mining, which can be used for mission optimization and maneuver prediction to assist air combat decision-making.

1. Introduction
In the network era, avionics system not only contains the massive data of land, sea, air, and space, but also has a variety of heterogeneous data, including early warning, reconnaissance, intelligence, confrontation, etc. With the explosive growth of information, the requirements for information processing are constantly improving. How to quickly and accurately integrate, relate a large number of scattered, isolated intelligence, transform information advantages into knowledge advantage, then transform into decision-making advantage, becomes the key to fight for air power, and also the problem that knowledge graph needs to solve.

It is necessary to construct a professional aviation knowledge graph to alleviate the problem of information overload, improve efficiency and provide more useful information for commanders. Starting from the architecture, this paper expounds the construction process and application examples of aviation knowledge graph.

2. Technical Architecture of Knowledge Graph
The technical architecture of knowledge graph contains the construction and application of knowledge graph, where the construction process including knowledge extraction, knowledge fusion, knowledge processing [1], knowledge update, etc., as shown in Figure 1. The construction of knowledge graph is based on the combination of natural language processing, text mining, machine learning, artificial methods and so on, which can be applied to complex relationship query, human-computer intelligent interaction and other directions in the aviation field.
2.1. Data Acquisition

The original data sources of knowledge graphs include structured data such as relational database, semi-structured data stored in document and web formats, and unstructured data stored in pictures, audio and video formats, etc. Therefore, making full use of these data and processing them correctly can be a good basis for the construction of knowledge graphs.

2.2. Knowledge Extraction

Knowledge extraction is an important step in the construction of knowledge graphs, it automatically extracts entity names, attributes, attribute values and relationships among entities from heterogeneous...
data to build domain ontology. The extraction methods include rule-based extraction [2], method based on machine learning [3] and open domain-oriented extraction method [4], etc.

Entity extraction is mainly for unstructured data such as image and text [5]. For image data, text information can be recognized by OCR technology, so that the image can be converted into text. The recognition process includes five main steps: image input, image pre-processing, feature extraction, classification and post-processing.

Relation extraction is mainly for unstructured data such as text. The associated relationship are extracted between entities from relevant corpus to solve the problem of semantic connection between discrete entities, also form a network of knowledge structure. In order to get rid of the limitation of traditional artificial method, machine learning method is introduced into the relation extraction. The common models of relation extraction [3,6] include convolution neural network, capsule neural network and other deep learning models, such as relation extraction model based on CNN, relation extraction model based on RNN and relation extraction model based on PCNN.

The purpose of attribute extraction is to collect the attribute information of specific entity from different sources, and realize the complete description of entity attribute. The Methods based on rules, heuristic algorithm, and the algorithms based on machine learning can be used for attribute extraction.

2.3. Knowledge Fusion
The purpose of knowledge fusion [7] is to eliminate the ambiguity of concepts, remove redundancy and misconceptions so as to ensure the quality of knowledge. Knowledge fusion includes two parts: entity linking and knowledge combination. Entity linking is for semi-structured data and unstructured data, and the knowledge combination is aimed at structured data.

2.3.1. Entity Linking. Entity linking make the extracted entity corresponds to the homologous entity in the knowledge base. The general process is:
1) Obtain the entity appellation through entity extraction;
2) Ensure correctness and uniqueness of the entity with entity disambiguation and coreference resolution;
3) Link the entity appellation to the corresponding entity in the knowledge base.

Entity disambiguation is used to solve ambiguity caused by the same-named appellation pointing to different entities. Clustering method is mainly used in entity disambiguation. There are four common models of clustering disambiguation: ①The model of Space vector; ②The semantic model; ③The model of Social network; ④The model of encyclopedic knowledge.

Coreference resolution is used to solve ambiguity caused by multiple appellations correspond to the same entity in the real world. Machine learning method is introduced to improve the accuracy of the algorithm, such as decision trees algorithm, Hobbs algorithm and centering theory.

2.3.2. Knowledge Combination. There are two main types of knowledge combination: ① The combination of external databases such as DBpedia [8], YAGO [9-10], Wikipedia [10], Freebase [11] and so on. This step mainly deals with two levels of problems: one is fusion in the data layer, including entity’s reference, attribute, relationship and category, etc. The main purpose is to avoid relationship conflicts in the instances which can result in redundancy; another one is to integrate the new ontology into the existing ontology database through fusion in the pattern layer. ② The combination of relational databases. An important source of high-quality knowledge is the domain’s own relational database. In order to integrate these structured historical data into the knowledge graph, resource description framework (RDF) [12], can be used as the data model. The essence is to convert the data of relational database into the triple data of RDF.

2.4. Knowledge Processing
Knowledge processing includes three aspects: ontology construction, knowledge inference and quality evaluation.
2.4.1. **Ontology Construction.** Ontology is a formal description of domain concepts. It contains the basic concepts and the relationship between concepts in the domain that can be interpreted by machines, and a general dictionary is defined for experts or machines in a specific domain to share the information. The process of ontology construction is as follows:

1. **Step 1:** Determine the domain and scope of ontology;
2. **Step 2:** Consider reusing the existing ontology;
3. **Step 3:** Enumerate important items in the ontology;
4. **Step 4:** Define the class and its hierarchy;
5. **Step 5:** Define the properties of the class, which called slot;
6. **Step 6:** Define constraints on slots;
7. **Step 7:** Create an instance.

2.4.2. **Knowledge Inference.** Knowledge inference [13] can mine hidden knowledge on the basis of existing knowledge base, obtain new knowledge or conclusions through computer inference, so as to meet the semantic requirements, expand and enrich the knowledge network. Its specific tasks can be divided into satisfiability, classification and materialization. Knowledge inference includes rule inference, factual inference, and distributed inference, etc.

2.4.3. **Quality Evaluation.** It is used to evaluate the credibility of new data and ensure the quality of knowledge. In view of the specialization of domain knowledge, it is necessary to introduce domain expert guidance.

2.5. **Knowledge Update.**

The update of knowledge graph mainly includes the update of schema layer and data layer.

The update of schema layer refers to the update of elements in ontology, including the addition, modification, deletion of concepts, the update of concept’s attribute and hyponymy. It usually requires manual intervention to verify the correctness and rationality of the updated knowledge through the test method based on feedback iteration. The update of data layer refers to the update of entity elements, including the addition, modification, deletion of entities, as well as the update of relationship and attribute values, which is realized automatically through the methods of fusion, similarity matching, association mapping and so on.

2.6. **Applications**

- **Accurate Search**
  In the network environment, information or situation is changing rapidly, and massive information is interconnected with complex multi-dimensional structure. By establishing such complex relationship, knowledge graph can not only search complex multi-source objects such as text, image, video, code, etc., but also realize the accurate search [14] for large data through sentence-level search and other methods, which can be used for precise classification, semantic understanding, accurate search for tactical intent, etc.

- **Efficient Query**
  Users can quickly and directly acquire relevant knowledge by compiling information requirements into query statements in a specific format. For example, the “automatic conversion method of query statement based on Rule Template matching” can be used.

- **Intelligent Human-Computer Interaction**
  The system can respond to pilot’s questions based on the information provided by knowledge graph, thus forming an intelligent question answering system [15], in which dialogue interaction replaces text-searching as the main interaction mode.

- **Intelligent Recommendation and Assistant Decision**
  With the help of knowledge graph, the situation information such as the flight status of target and battlefield environment is perceived, the deep relationship is inferred by mining the implicit
relationship, and the intelligent recommendation is given by combining current situation information in the environment, so as to realize the function of assistant decision.

3. Aviation knowledge graph
   When complete and accurate environmental information is available in advance, the optimal flight mission can be planned at one time. The actual situation is that it is difficult to ensure that the acquired environmental information does not change; on the other hand, due to the uncertainty of the mission, it is often necessary to temporarily change the flight mission. Knowledge graph can help to make better use of the coordination between domain experts and computer systems, and realize the optimal mission planning through resource sharing and information mining.

Figure 2. An instance of knowledge graph.

Figure 2 is an instance of the knowledge graph in aviation, for the air combat decision-making, which directly shows the cross-linking relationship among the types of aircraft, task, maneuver, equipment and weapons, as well as related vehicles, helicopters, warships, other objects and their attributes for air combat. With the help of knowledge graph, the hidden relationship between task and maneuver can be mined for maneuver prediction and assistant decision-making.
4. Conclusion

Information explosion brings a large amount of data overload in aviation, which greatly reduces the utilization rate of effective information and affects the commander to make correct decisions. It is vital to construct the aviation knowledge graph. This paper expounds the technical architecture of knowledge graph, and constructs an instance of knowledge graph for air combat, which can support flight planning, maneuver prediction, assistant decision and other application scenarios.

References

[1] Nadeau D, Sekine S. A survey of named entity recognition and classification[J]. Lingvisticae Investigationes, 2007, 30(1) pp 3-26.
[2] Reiss F, Raghavan S and Krishnamurthy, et al. An Algebraic Approach to Rule-Based Information Extraction[C]. IEEE International Conference on Data Engineering, 2008, pp 1-10.
[3] Zeng D, Liu K, Lai S, et al. Relation Classification via Convolutional Deep Neural Network[C]// COLING. 2014: pp 2335-2344.
[4] A. Fader, S. Soderland and O. Etzioni. Identifying relations for open information extraction[C]. Proceedings of Conference on Empirical Methods in Natural Language Processing. EMNLP, 2011, pp1535-1545.
[5] L.F. RAU. Extracting company names from text[C]. Proc of the 7th IEEE conference on Artificial Intelligence Applications. NJ: IEEE,1991, pp 29-32.
[6] Bordes A, Glorot X, Weston J, et al. A semantic matching energy function for learning with multi-relational data[J]. Machine Learning, 2014, 94(2) pp 233-259.
[7] Dong X, Gabrilovich E, Heitz G, et al. Knowledge vault: A web-scale approach to probabilistic knowledge fusion[C]. Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2014, pp 601-610.
[8] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, and Z. Ives. DBpedia: A nucleus for a web of open data [C]. Proceedings of the 6th International Semantic Web Conference, Busan: springer, 2007, pp 722-735.
[9] F. M. Suchanek, G. Kasneci, and G. Weikum. YAGO: A core of semantic knowledge unifying wordnet and Wikipedia. Proceedings of the 16th International Conference on World Wide Web, Alberta: ACM, 2007, pp 697-706.
[10] J Hoffart, F. M. Suchanek, K. Berberich, YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia, 2013, International Joint Conference on Artificial Intelligence. 2007:pp 1962-1969.
[11] K. Bollacker, C. Evans, et al. Freebase: a collaboratively created graph database for structuring human knowledge. Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data, SIGMOD 2008, pp 1247-1250.
[12] C. Unger, L. Buhmann and J. Lehmann, et al. Template-based question answering over RDF data. ACM Proc. www, 2012, pp 639-648.
[13] R. Socher, D. Chen, C. D. Manning. Reasoning with neural tensor networks for knowledge base completion.[C]. International Conference on Neural Information Processing Systems. 2013, pp 926-934.
[14] Jiang K, Qian K, et al. Military Information Retrieval Technology Architecture Based on Knowledge Graph. Command Information System and Technology. 2016,7(1), pp 48-52.
[15] O. Kolomiyets, M.-F. Moens, A survey on question answering technology from an information retrieval perspective. Elsevier Trans. Inf. Sci. 181(24), 2011, pp 5412-5434.