International Journal of Engineering & Technology, 7 (3.33) (2018) 168-173

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Design of Link Evaluation Method to Improve Reliability based on Linked Open Big Data and Natural Language Processing

Yonglak SHON[1], Jaeyoung PARK[2], Jangmook KANG[3], Sangwon LEE[4]*

[1]*Department of Computer Engineering, Seokyeong University, Seoul, Korea*
[2]*Department of Computer Engineering, Seokyeong University, Seoul, Korea*
[3]*Department of Big Data & Industry Security, Namseoul University, Cheonan, Korea*
[4]*Department of Computer & Software Engineering (Institute of Convergence Creativity), Wonkwang University, Iksan, Korea*
**Corresponding author E-mail: sangwonlee@wku.ac.kr*

Abstract

The LOD data sets consist of RDF Triples based on the Ontology, a specification of existing facts, and by linking them to previously disclosed knowledge based on linked data principles. These structured LOD clouds form a large global data network, which provides a more accurate foundation for users to deliver the desired information. However, it is difficult to identify that, if the presence of the same object is identified differently across several LOD data sets, they are inherently identical. This is because objects with different URIs in the LOD datasets must be different and they must be closely examined for similarities in order to judge them as identical. The aim of this study is that the prosed model, RILE, evaluates similarity by comparing object values of existing specified predicates. After performing experiments with our model, we could check the improvement of the confidence level of the connection by extracting the link value.

Keywords: *AI; Big Data; Linked Open Data; NLP; Ontology; Ontology Web Language; Resource Description Framework*

1. Introduction

Today's Web is rapidly evolving from the Web of documents to the Web of data [1, 2, 9, 12]. The Web of documents publishes a generally structured document unit, and the Web of data publishes a data unit consisting of existing individual facts. The web of data allows users to actively extend their knowledge in the future by actively linking newly released facts (data) to previously released facts (data). This paper presents a more reliable means of making these connections.

Documents that are released on the Web of a document provide high legibility, but make automatic and precise access through applications difficult. In order to solve the problem of having multiple data within a document not individually identified by the application but identified by a single document, the Linking Open Data (LOD) [3] emerged under the W3C initiative.

LOD is the free disclosure of data sets on the web in accordance with the Linked Data principle based on an open license. LOD provides high openness. In recent years, the transition to LOD forms consisting of Resource Description Framework (RDF) [13, 15] triples is actively underway to make public data more generally available.

The LOD Cloud, the web of data, comprises a set of data released in the form of LOD, and individual facts in each data set are described as RDF Triple as shown in Figure 1. LOD datasets provide existing facts in RDF Triple structures in the form of subject, predicate, and object. This enables semantic [4, 14] processing of facts, further linking with other facts and reasoning through the extension of these connections.

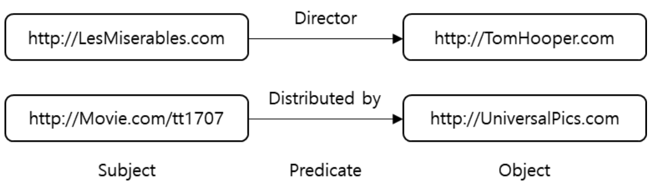

Fig. 1: RDF Triple

LOD Cloud's openness is faithful to the concept of 'Anyone can say Anything about Anything' advocated by Tim Berners Lee [3]. In addition, through LOD datasets built under these slogans, knowledge is expanded. In order for this wisdom to ultimately become an expression of wisdom, there must be a way to connect each triple with each other. LOD is based on the Ontology [10, 11] which is a statement of existing facts. The ontology requires that the objects that make up the facts be uniquely identifiable. For this purpose, LOD uses the Uniform Resource Identifier (URI). If two sub-states appearing in Figure 1 are judged to be the same, connect their URI as shown in Figure 2. By describing as RDF Triple, knowledge can be expanded through given facts.

<http://LesMiserables.com> <owl:sameAs> <http://Movie.com/tt1707>

Fig. 2: Object Linkage

2. Related Works

In LOD clouds, the method of connecting triple to triple is largely divided into passive and automatic methods. The passive method is for people to read, understand, and connect directly to each of the triples, with good connection quality but not to handle large triples. In contrast, automatic methods by computer systems may degrade connection quality relatively more than the passive method, but are not affected by the triple scale. This paper proposes to allow you to reconsider the quality of the connection while connecting the triple in an automatic manner.

SILK and LIMES are the typical ways in which a computer system automatically connects the triple. They have adopted a method of comparing object values of specified predicates in connection policies adopted by the public to find out how much their subjects match. However, simply comparing the object values can lead to situations where these objects are not sufficiently similar, even though they are actually identical. The following example presents the problems of the existing similarity assessment methods.

This paper presents a Reliability Improving Link Evaluation (RILE) as a way to provide a more reliable connection between the triple by reducing these problems. The RILE evaluates a new fact derived through reasoning in addition to considering the grammatical characteristics of the predicates in addition to the existing ObLE (Object Based Link Evaluation) method.

RILE applies to LOD datasets as RDFS (RDF Schema) and OWL 2. RDFS and OWL 2 are the languages that define a strict vocabulary structure that describes RDF triple and enables reasoning. The reason why they were selected is that the W3C [6, 7, 8] actively supports them, and the languages such as semantic Net, frame, and technology description (most of them). The vocabulary that appears in many LOD datasets that make up the LOD cloud is limited to RDFS. However, data sets using the vocabulary defined by OWL 2 are also steadily increasing. Considering the nature of the grammatical elements that RDFS and OWL 2 have in this trend, the RILE will be able to help improve the LOD cloud and facilitate the use of LOD datasets.

A method such as 'comparing object values of a given term' has limits to situations in which an object is assessed as not sufficiently similar, even though it is actually identical. If there is actually a difference in object values between the same objects because the language described is different or the time written, the existing approach is unable to identify and give a similar level of trust in the cloud. That is, the similarity between objects cannot be determined simply by comparing the object values, and the similarities between the objects, such as structural features, constraints, and properties, must be considered. Thus, in addition to comparing object values of specified predicates, this paper provides a closer look at the grammatical elements of the terms RDFS / OWL 2 to identify and base the weights on them. In addition, new facts derived from reasoning based on specific grammatical factors are reflected in the similarity assessment to minimize the occurrence of such a situation, thereby minimizing the reliability of the LOD cloud.

3. Reliability Improving Link Evaluation

3.1. Concept of RILE

In evaluating the equality of objects described in different triples, RILE is derived through reasoning, along with the OBLE method as well as the grammatical characteristics of the triple-letter parts. Through this, we expect to further improve the level of confidence in the assessment of similarity or difference between the triplets. The LOD data set is made up of a large number of triples, each of which consists of [subjects, predicates, objects]. As illustrated in Figure 1, the individual facts described in triple units can be extended to one another as long as they are practically identical to the other. This means that it is possible to express the potentially existing fact through these connections as if it were newly real. However, it is not possible under any circumstances to ensure full confidence in the newly established facts with the facts explicitly disclosed. It is also not possible for the system to quantify completely and automatically the level of trust that the connection has, except in completely same cases, or in completely separate cases. However, while the LOD framework presented by W3C is constantly changing, it is the highest value sought by information disclosure, as it always shows at the top that a confidence layer is always located. In response, RILE tries to improve the level of connectedness between the facts they express through careful examination of the grammatical elements that triple has.

RILE proceeds to evaluate the similarity among the triple at a certain level by applying an inference, with input triples similar to the existing OBLE method. Follow the steps to consider the grammatical elements of each triple with respect to the results. This two-phase evaluation structure ensures that the results of the connectivity assessment between the triple peers are more reliably improved than the traditional methods. Assessment results of the RILE are not necessarily similar. Traditional methods may evaluate similar facts that may never be the same in some cases. In this case, the review of grammatical elements in the second step can show that they can never be identical.

RILE is a descriptive text for grammar. Consider by dividing it into subject-predicate and lexical grammatical elements. The grammar factors considered correspond to a triple predicate part. Predicate part techniques may exist as a subject-predicate part of the LOD data set, on the other hand as a lexical part. The LOD data sets consist of RDF Triple and each RDF Triple is described in a structure describing the subject in RDFS and OWL2 grammar.

Table 1: The Effect of Similarity Evaluation on the Subject-Predicate Grammar Elements

Type	Grammar Elements	Effect
RDFS	rdfs:Resource, rdfs:Class, rdfs:Literal, rdfs:Datatype, rdf:XMLLiteral, rdfs:isDefinedBy	No
OWL	owl:Class, owl:Thing, owl:Nothing	
OWL2	owl:Annotation, owl:OnClass, owl:NameIndividual, owl:deprecated	
RDFS	rdfs:subClassOf, rdfs:seeAlso, rdfs:member, rdfs:label, rdfs:comment	Comparative
OWL	owl:oneOf, owl:intersectionOf, owl:equivalentClass, owl:unionOf	
OWL2	owl:disjointUnionOf, owl:members	
RDFS		Absolute
OWL	owl:sameAs, owl:complementOf, owl:disjointWith, owl:differentFrom, owl:distinctMembers, owl:AllDifferent	
OWL2	owl:AllDisjointClasses, owl:hasKey	

A review of subject-predicate grammatical elements performed by RILE corresponds to the consideration in this structural dimension of RDF Trips. On the other hand, the other important roles of RDFS and OWL 2 are to provide grammar for defining the vocabulary used by data producers to describe RDF Triple on a topical basis. Vocabulary is an essential tool to understand, share, communicate and utilize LOD datasets. In order to define these words, grammar must be provided, and RDFS and OWL 2 play their part. Therefore, RILE considers the grammatical elements of the words used on RDF Triple and actively utilizes the results to evaluate similarity.

Table 2: The Effect of Similarity Evaluation on the Lexical Grammatical Elements

Type	Grammar Elements	Effect
RDFS	rdf:Property	No
OWL	owl:Restriction, owl:onProperty, owl:ObjectProperty, owl:DatatypeProperty, owl:OntologyProperty, owl:imports, owl:versionInfo, owl:priorVersion, owl:backwardCompatibleWith, owl:incompatibleWith, owl:DeprecatedClass, owl:DeprecatedProperty	

OWL2	owl:annotatedProperty, owl:annotatedSource, owl:annotatedTarget, owl:assertionProperty, owl:bottomDataProperty, owl:DataRange, owl:bottomObjectProperty, owl:Axiom, owl:datatypeComplementOf, owl:onDataRange, owl:targetIndividual, owl:topDataProperty, owl:topObjectProperty, owl:withRestrictions, owl:versionIRI, owl:AnnotationProperty, owl:propertyDisjointWith, owl:targetValue, owl:AlldisjointProperty, owl:AsymmetricProperty, owl:onDatatype, owl:sourceIndividual	
RDFS	rdfs:domain, rdfs:range, rdfs:subPropertyOf	Comparative
OWL	owl:allValuesFrom, owl:someValuesFrom, owl:hasValue, owl:inverseOf, owl:maxCardinality, owl:minCardinality, owl:equivalentProperty, owl:TransitiveProperty, owl:SymmetricProperty, owl:cardinality	
OWL2	owl:IrreflexiveProperty, owl:ReflexiveProperty, owl:qualifiedCardinality, owl:propertyChainAxiom, owl:NegativePropertyAssertion, owl:hasSelf, owl:minQualifiedCardinality, owl:maxQualifiedCardinality	
RDFS		Absolute
OWL	owl:FuntionalProperty, owl:InverseFuntionalProperty	
OWL2		

In addition to the consideration of grammatical elements, it is possible for RILE to obtain an improved level of similarity by reflecting the reasoning in its assessment. At LOD, reasoning enables the expression of new facts. One fact consists of the relationship between the data and the data, and these facts combine to construct LOD. The reasoning leads to new relationships based on the data and relationships stated in the existing facts that make up the LOD, which lead to the expression of new facts. Thus, reasoning is the ultimate goal of Ontology. Taking this into account, RILE can achieve a more reliable level of similarity by reflecting new facts derived from the reasoning of specific grammatical elements in the evaluation of similarity.

In order to understand the impact of the grammatical factors under consideration by RILE, this chapter examines some examples under the theme of ‘Les Miserables’. Les Miserables revolves around the main character Jean Valjean, the story of many poor French people living around 1830. The novel, Les Miserables, gained much popularity through its literary value as well as its historical value, its historical perspective, life style, and other historical value. If these same themes are made with different media, each medium is likely to have different facts because different media have different settings or descriptions. Les Miserables will be able to effectively explain these characteristics. RILE considers all the grammatical elements listed in Table 1 and Table 2, but only a few situations are given as examples. In the examples below, triple stars appear from side to side. The left is part of the newly released source LOD data set and the right is part of the targeted LOD data set to which it is to be connected.

3.2. Evaluation on the Subject-Predicate Grammar Elements

The predicate-predicate descriptive grammar element that the predicate possesses allows more sophisticated descriptions of existing facts through the structuralization of the triple. In other words, the definition of a subject or set of objects and a hierarchy is possible, and it plays an effective role in identifying identical or non-uniform objects based on it. For example, among the grammatical elements, ‘rdfs : subclass’ defines the hierarchy of objects. ‘rds:seeAlso’ is used to link another fact associated with the subject, and ‘owl:sames’ are elements of the equality of objects. Consider these factors to explore the process of incorporating them into the assessment of similarity. The examples below describe the fact that Jean Valjean, the character of the movie Les Miserables, and Jean Valjean, the character of the novel Les Miserables. Both Jean Valjean are the same character and the facts are being told slightly differently depending on the media.

(Example 1) Similarity Evaluation Considering Hierarchy (Figure 3): The following example illustrate how similar levels of similarity among objects can be helped if other objects associated with the objects are the same.

```
<http://LesMiserables.com/Char/JeanValjean>
  <rdfs:label>              <JeanValjean>
  <Les:PNumber>             <24601>
          <rdfs:subClassOf>
   <http://LesMiserables.com/Char>
-------------------------------------------
<http://VictorHugo.com/Les/JeanValjean>
  <rdfs:label>              <Madeleine>
  <Les:PNumber>             <24601>
          <rdfs:subClassOf>
   <http://LesMiserables.com/Char>
```

Fig. 3: rdfs:subClassOf

(Example 2) Similarity Evaluation Considering Related Objects (Figure 4): The following example illustrate how similar levels of similarity among objects can be helped if other objects associated with the objects are the same.

```
<http://LesMiserables.com/Char/JeanValjean>
  <rdfs:label>              <JeanValjean>
  <Les:PNumber>             <24601>
          <rdfs:seeAlso>
   <http://LesMiserables.wikia.com>
-------------------------------------------
<http://VictorHugo.com/Les/JeanValjean>
  <rdfs:label>              <Madeleine>
  <Les:PNumber>             <24601>
          <rdfs:seeAlso>
   <http://LesMiserables.wikia.com>
```

Fig. 4: rdfs:seeAlso

(Example 3) Similarity Evaluation Considering Comment (Figure 5): The following example illustrate how similar levels of objects can be assessed better if they have the same description.

```
<http://LesMiserables.com/Char/JeanValjean>
  <rdfs:label>              <JeanValjean>
  <Les:PNumber>             <24601>
          <rdfs:comment>
   "The LesMiserables Characters"
-------------------------------------------
<http://VictorHugo.com/Les/JeanValjean>
  <rdfs:label>              <Madeleine>
  <Les:PNumber>             <24601>
          <rdfs:comment>
   "The LesMiserables Characters"
```

Fig. 5: rdfs:comment

(Example 4) Similarity Evaluation Considering Object Identity (Figure 6): The following example illustrates when two objects determine their identity by identifying grammatical elements that mean they are identical. This case governs the results of a similar level of assessment of the other facts of the object.

```
<http://LesMiserables.com/Char/JeanValjean>
  <rdfs:label>              <JeanValjean>
  <Les:PNumber>             <24601>
          <owl:sameAs>
   <http://LesM.wiki.com/JeanValjean>
-------------------------------------------
<http://VictorHugo.com/Les/JeanValjean>
  <rdfs:label>              <Madeleine>
  <Les:PNumber>             <24601>
          <owl:sameAs>
   <http://LesM.wiki.com/JeanValjean>
```

Fig. 6: owl:sameAs

(Example 5) Similarity Evaluation Considering Identity of Objects to Which They Belong (Figure 7): The following example illustrates that objects to which two objects belong can be highly evaluated if they are identical to a particular object.

```
<http://LesMiserables.com/Char/JeanValjean>
  <rdfs:label>                    <JeanValjean>
  <rdf:type> <http://LesM.com/Characters>
       <http://LesM.com/Characters>
  <owl:equivalentClass> <http://L.co/C>
-------------------------------------------
<http://VictorHugo.com/Les/JeanValjean>
  <rdfs:label>                    <Madeleine>
  <rdf:type> <http://VHugo.com/L/Char>
       <http://VHugo.com/L/Char>
  <owl:equivalentClass> <http://L.co/C>
```

Fig. 7: owl:equivalentClass

(Example 6) Similarity Evaluation Considering Identity of Objects to Which They Belong (Figure 8): The following example illustrates that the objects to which two objects belong are intersection sets of the same objects can be reflected in the similarity assessment, resulting in a higher level of similarity.

```
<http://LesMiserables.com/Char/JeanValjean>
  <rdfs:label>                    <JeanValjean>
  <rdf:type> <http://LesM.com/Characters>
       <http://LesM.com/Characters>
  <owl:intersectionOf> <(Person, Char)>
-------------------------------------------
<http://VictorHugo.com/Les/JeanValjean>
  <rdfs:label>                    <Madeleine>
  <rdf:type> <http://VHugo.com/L/Char>
       <http://VHugo.com/L/Char>
  <owl:intersectionOf> <(Person, Char)>
```

Fig. 8: owl:intersectionOf

(Example 7) Similarity Evaluation Considering Union of Objects to Which They Belong (Figure 9): The following example draws a better level of similarity in which these factors are taken into account when both objects are composed of a combination of the same object.

```
<http://LesMiserables.com/Char/JeanValjean>
  <rdfs:label>                    <JeanValjean>
  <rdf:type> <http://LesM.com/Characters>
       <http://LesM.com/Characters>
  <owl:unionOf> <(LesMiserC, Person)>
-------------------------------------------
<http://VictorHugo.com/Les/JeanValjean>
  <rdfs:label>                    <Madeleine>
  <rdf:type> <http://VHugo.com/L/Char>
<http://VHugo.com/L/Char>
  <owl:unionOf> <(LesMiserC, Person)>
```

Fig. 9: owl:unionOf

(Example 8) Similarity Evaluation Considering Description of Object Difference (Figure 10): The following example illustrates when two objects determine their differences by identifying grammatical elements that are different. This case governs the results of a similar level of assessment of the other facts of the object.

```
<http://LesMiserables.com/Char/JeanValjean>
  <rdfs:label>                    <JeanValjean>
          <owl:differentFrom>
   <http://MrPizza.co/Menu/JeanValjean>
-------------------------------------------
<http://MrPizza.co/Menu/JeanValjean>
  <rdfs:label> <JeanValjean>
```

Fig. 10: owl:differentFrom

(Example 9) Similarity Evaluation Considering Combined Nature of Objects (Mutually Beta) that the Objects Belong to (Figure 11): The following example illustrates a level of similarity in the evaluation of similarity if the objects belonging to both populations consist of a combination of the same objects (mutually exclusive).

```
<http://LesMiserables.com/Char/JeanValjean>
  <rdfs:label>                    <JeanValjean>
  <rdf:type> <http://LesM.com/Characters>
       <http://LesM.com/Characters>
  <owl:disjointUnionOf> <(Human,Char)>
-------------------------------------------
<http://VictorHugo.com/Les/JeanValjean>
  <rdfs:label>                    <JeanValjean>
  <rdf:type> <http://VHugo.com/L/Char>
<http://VHugo.com/L/Char>
  <owl:disjointUnionOf> <(Human,Char)>
```

Fig. 11: owl:disjointUnionOf

(Example 10) Similarity Evaluation Considering Identification of Objects to Which They Belong (Figure 12): The following example describes when an object to which both objects belong is uniquely identifiable by the object value of a particular predicate, when the object value is the same, the object must be the same. These cases ensure the identity of the objects, whether or not they match the other facts of the object.

```
<http://LesMiserables.com/Char/JeanValjean>
  <rdfs:label>                    <JeanValjean>
  <Les:PNumber>                   <24601>
  <rdf:type> <http://LesM.com/Characters>
       <http://LesM.com/Characters>
  <owl:hasKey> <Les:PNumber>
-------------------------------------------
<http://VictorHugo.com/Les/JeanValjean>
  <rdfs:label>                    <Madeleine>
  <Les:PNumber>                   <24601>
  <rdf:type> <http://LesM.com/Characters>
<http://LesM.com/Characters>
  <owl:hasKey> <Les:PNumber>
```

Fig. 12: owl:hasKey

3.3. Evaluation on the Lexical Grammatical Elements

The lexical grammar element of a prediate is a factor that indicates the nature of a vocabulary, which allows a more reliable evaluation of the level of similarity. This means that the nature of objects through the logical nature of words is identified and a similarity assessment is conducted based on them to improve the reliability of the connection. Look at how these lexical grammatical elements can affect the evaluation of the similarities between individuals.

(Example 11) Similarity Evaluation Considering Inverse Function Property (Figure 13): The following example shows the process of using these properties to determine the equality and disparity of objects if a predicate with functional dependencies is used in assessing similarity.

```
<http://LesMiserables.com/C/Montreuil>
<rdfs:label>          <Montreuil        City>
<Les:Code> "74"
-------------------------------------------
<http://VictorHugo.com/Loc/Montreuil>
<rdfs:label>                      <Location>
<Les:Code> "74"
-------------------------------------------
<Les:Code>
  <rdf:type> <owl:InverseFunctionalProperty>
```

Fig. 13: owl:InverseFunctionalProperty

(Example 12) Similarity Evaluation Considering Domain of Subject-Object (Figure 14): The following example presents a situation in which the results of an existing assessment can be further improved if the similarity assessment between the two objects has the same meaning of the different words.

```
<http://LesMiserables.com/C/Montreuil>
  <rdfs:label>        <Montreuil        City>
  <Les:mayor> <Madeleine>
-------------------------------------------
<Les:mayor>
  <rdfs:domain> <http://L.com/LOC>

<http://VictorHugo.com/Loc/Montreuil>
  <rdfs:label>                    <Location>
  <VH:mayor> <Madeleine>
-------------------------------------------
<VH:mayor>
  <rdfs:domain> <http://L.com/LOC>
```

Fig. 14: rdfs:domain

(Example 13) Similarity Evaluation Considering Cardinality of Vocabulary (Figure 15): In the following example, the two predicates may be more likely to be the same predicate if the two terms used in the assessment of similarity are the same, thus improving the results of the existing assessment.

```
<http://LesMiserables.com/C/Montreuil>
<rdfs:label>        <Montreuil        City>
<Les:mayor> <Madeleine>

<owl:Restriction>
  <owl:onProperty> <Les:mayor>
  <owl:cardinality> xsd:decimal "1"
```

```
<http://VictorHugo.com/Loc/Montreuil>
<rdfs:label>                    <Location>
<VH:mayor> <Madeleine>

<owl:Restriction>
  <owl:onProperty> <VH:mayor>
  <owl:cardinality> xsd:decimal "1"
```

Fig. 15: owl:cardinality

(Example 14) Similarity Evaluation Considering Maximum Cardinality of Vocabulary (Figure 16): The following example shows that two predicates are more likely to be predicate if the maximum number of two predicates used in assessing similarity

```
<http://LesMiserables.com/C/Montreuil>
<rdfs:label>        <Montreuil        City>
<Les:mayor> <Madeleine>

<owl:Restriction>
  <owl:onProperty> <Les:mayor>
  <owl:maxCardinality> xsd:decimal "3"
```

```
<http://VictorHugo.com/Loc/Montreuil>
<rdfs:label>                    <Location>
<VH:mayor> <Madeleine>

<owl:Restriction>
  <owl:onProperty> <VH:mayor>
  <owl:maxCardinality> xsd:decimal "3"
```

Fig. 16: owl:maxCardinality

(Example) Similarity Evaluation Considering Predicate Hierarchy (Figure 15): The following example shows that the two predicates used in the evaluation of similarity are likely to be the same predicate if they have the same hierarchy.

```
<http://LesMiserables.com/C/Montreuil>
<rdfs:label>        <Montreuil        City>
<Les:mayor> <Madeleine>

<Les:mayor>
  <rdfs:subPropertyOf> <City:Council>
```

```
<http://VictorHugo.com/Loc/Montreuil>
<rdfs:label>                    <Location>
<VH:mayor> <Madeleine>

<VH:mayor>
  <rdfs:subPropertyOf> <City:Council>
```

Fig. 17: rdfs:subPropertyOf

(Example) Similarity Evaluation Considering Predicate Equivalence (Figure 18): The following examples can confirm that different predicates are the same prior if they are the same as a particular predicate, based on which the similarity of the object can be made.

```
<http://LesMiserables.com/C/Montreuil>
<rdfs:label>        <Montreuil        City>
<Les:mayor> <Madeleine>

<Matrix:Commander>
  <owl:equivalentProperty> <City:Voc>
```

```
<http://VictorHugo.com/Loc/Montreuil>
<rdfs:label>                    <Location>
<VH:mayor> <Madeleine>

<MaComic:Commander>
  <owl:equivalentProperty> <City:Voc>
```

Fig. 18: owl:equivalentProperty

(Example 16) Similarity Evaluation Considering Predicate Constraint (Figure 19): The following example shows that if the predicates have the same specific constraints in different situations, the same predicate probability increases and as a result affects the assessment of similarity of objects.

```
<http://LesMiserables.com/C/Montreuil>
<rdfs:label>        <Montreuil        City>
<Les:mayor> <Madeleine>

<owl:Restriction>
  <rdf:type> <owl:NegativePropertyAssert
    ion>
  <owl:assertionProperty> <Les:mayor>
  <owl:sourceIndividual> <Javert>
  <owl:targetIndividual> <Montreuil>
```

```
<http://VictorHugo.com/Loc/Montreuil>
<rdfs:label>                    <Location>
<VH:mayor> <Madeleine>

<owl:Restriction>
  <rdf:type> <owl:NegativePropertyAsser
    tion>
  <owl:assertionProperty> <VH:mayor>
  <owl:sourceIndividual> <Javert>
  <owl:targetIndividual> <Montreuil>
```

Fig. 19: NegativePropertyAssertion

3.4. RILE Algorithm

RILE enhances and aggregates existing measurement values by reflecting the grammatical elements of tripleing after a similarity assessment of the OBLE method applied to the inference system. In addition, the similar values for the connection are analyzed and the results are notified to disconnects to determine the adequacy of the results of the similarity assessment. Since then, the connection is finally confirmed, enabling the enhancement of the level of participation of provider and confidence in the connection. The sequence of progression of the RILE integrated algorithm shown below is as follows (Figure 20).

```
INPUT SourceTriples ST, TargetTriples TT, LinkSpecification LS, Syntax LinkSpecification SLS;
OUTPUT LinkDecisionTripleSet LDTS;
BEGIN SIMILARITY_EVALUATION
  foreach ST.Triple_i ∈ ST.Triples do
    foreach TT.Triple_j ∈ TT.Triples do
      SimilarSet SS ← Similarity_Compare(ST.Triple_i, TT.Triple_j, LS, TT);
    end
  end
  foreach SS.Triple_i ∈ SS.Triples do
    foreach SLS.Predicate_j ∈ SLS.Predicates do
      if (SLS.Predicate_j = SS.Triple_i.CompareResult.Predicate) then
        SS.Triple_i.Score *= SLS.Predicate_j.weight;
      end if
    end
  end
  SLS.Aggregate(SS.Triples);
  SyntaxAddedTripleSet SATS, LinkDecisionTripleSet LDTS;
  foreach SS.Triple_i ∈ SS.Triples do
    if (SS.Triple_i.AggreScore >= SLS.Threshold) then
      SATS.AssertTriple(SS.Triple_i);
    end if
  end
  LDTS ← Link_Evaluation_Reporting_System(SATS);
END SIMILARITY_EVALUATION
```

Fig. 20: RILE Algorithm

4. Experiments

The performance comparison of RILE and OBLE for the same data set has shown that RILE has achieved better connectivity reliability than in conventional methods. For this purpose, the implementation and experimental environment of RILE and OBLE are as follows: (1) Hardware: Intel ® Core ™ i5/OS 3.40 GHz (4 CPUs) (2) Software: C #, dotNetRDF (1.0.0.) [16], Visual Studio 2010 Professional (.NET Framework 4.0) (3 Dataset: Same sub-object pair (500 pairs, 2,000 triple), different sub-order pairs (500 pairs, 2,000 triple)

The weight of grammatical elements specified by the publisher of the source LOD data set over the RILE was given a value between 0.15 and 0.33, and a confidence reference value of 0.8. The reason for weighting of these values for a particular interval is that it minimizes connection reliability distortions that can occur in the process of improving the measurement of similarity. Both RILE and OBLE also used ‘Levenshtein’ and ‘Average’ numbers, which were aggregated. The experiment was conducted on LOD datasets by comparing the results of connection evaluations derived from OBLE and RILE implementations. In addition, we confirm the change in similarity assessment according to the operation of the RILE inference system, how the weights affect the reliability of the coupling evaluation, and also provide the RILE for self-adjusting the weights. Figure 21 ~ 23 shows the results of the experiment from the viewpoint of evaluation on the subject predicate grammar elements. Figure 24~ 26 shows the results of the experiment from the viewpoint of evaluation on the lexical grammatical elements.

Fig. 21: Same Subjects’ Paris (Evaluation on the Subject Predicate Grammar Elements)

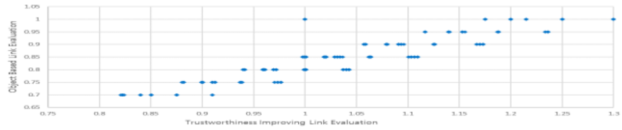

Fig. 22: Reliability Improving Link Evaluation (Evaluation on the Subject Predicate Grammar Elements)

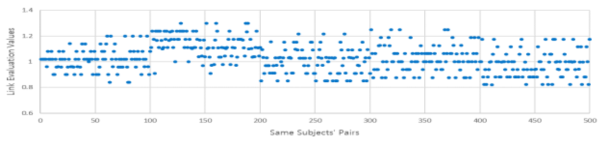

Fig. 23: Same Subjects’ Paris (Evaluation on the Subject Predicate Grammar Elements)

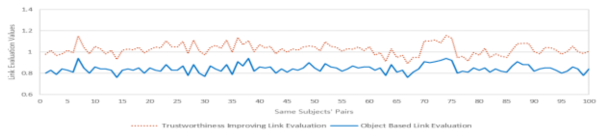

Fig. 24: Same Subjects’ Paris (Evaluation on the Lexical Grammatical Elements)

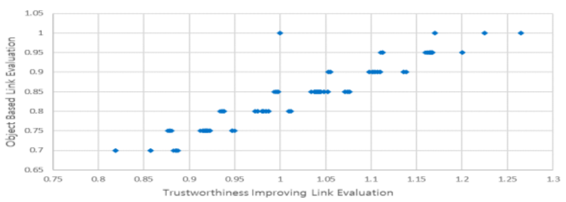

Fig. 25: Reliability Improving Link Evaluation (Evaluation on the Lexical Grammatical Elements)

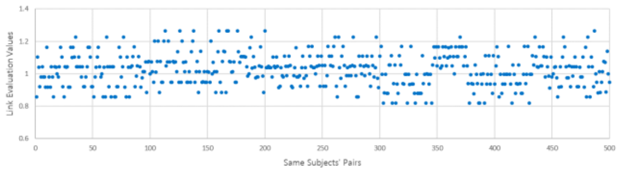

Fig. 23: Same Subjects’ Paris (Evaluation on the Lexical Grammatical Elements)

5. Conclusions

The RILE suggested in this paper evaluated similarity by comparing object values of existing specified predicates. The model introduced a system of reasoning, incorporating new facts derived from reasoning into the evaluation of similarity. In addition, by carefully considering the grammatical meaning of RDFS / OWL 2 of the predicates, this model produces a connection-level value by conducting a complementary assessment through weighting. Then, by realigning the weights and reassessing the connectivity based on the values of the derived connection values, the confidence level of the connection could be improved again.

Acknowledgement

This paper was supported by Wonkwang University in 2018.

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government(MSIP). (No.2018-0-00705, Algorithm design and software modeling for judge fake news based on artificial intelligence)

References

[1] J. Volz, C. Bizer, M. Gaedke, G. Kobilarov, “Silk – A Link Discovery Framework for the Web of Data”, Proc. of the 2nd Workshop on Linked Data on the Web 2009 (LDOW2009), 2009.

[2] A.-C. Ngonga Ngomo, S. Auer, “LIMES - A Time-Efficient Approach for Large-Scale Link Discovery on the Web of Data”, Proc. of the 22nd IJCAI, pp.2312-2317, 2011.

[3] SWEO Community Project. (2007, January 29). Linking Open Data [Online]. Available: http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData (accessed 2 November 2013).

[4] C. Bizer, T. Heath, T. Berners-Lee, “Linked data – the story so far”, International Journal on Semantic Web and Information Systems, Vol. 5, No. 3, pp.1-22, 2009.

[5] T. Berners-Lee. (2006, July 27). Linked Data [Online]. Available: http://www.w3.org/DesignIssues/LinkedData.html (accessed 2 November 2013).

[6] W3C. (1997, Agust 1). Resource Description Framework [Online]. Available: http://www.w3.org/RDF (accessed 2 November 2013).

[7] W3C. (1998, April 9). Resource Description Framework Schemas [Online]. Available: http://www.w3.org/TR/1998/WD-rdf-schema-19980409 (accessed 2 November 2013).

[8] W3C. (2008, April 11). OWL 2 Web Ontology Language: Primer [Online]. Available: http://www.w3.org/TR/2008/WD-owl2-primer-20080411 (accessed 2 November 2013).

[9] C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, S. Hellmann, “DBpedia – A Crystallization Point for the Web of Data”, Journal of Web Semantics, No. 7, pp.154-165, 2009.

[10] T.-R. Gruber, “A Translation Approach to Portable Ontology Specification”, Knowledge Acquisition, Vol. 5, No. 2, pp.199-200, 1993.

[11] W3C. (2002, July 29). OWL Web Ontology Language 1.0 Reference [Online]. Available: http://www.w3.org/TR/2002/WD-owl-ref-20020729 (accessed 2 November 2013).

[12] T. Heath, C. Bizer, Linked Data: Evolving the Web into a Global Data Space, 1st Ed., Morgan & Claypool, 2011.

[13] F. Scharffe, Y. Liu, C. Zhou, "RDF-AI: An Architecture for RDF Datasets Matching, Fusion, and Interlink", Proc. of the IJCAI 2009 Workshop on Identity, Reference, and Knowledge Representation (IR-KR), 2009.

[14] Y. Raimond, C. Sutton, M. Sandler, “Automatic Interlinking of Music Datasets on the Sema-ntic Web”, Proc. of the 1st Workshop about Linked Data on the Web 2008 (LDOW2008), 2008.

[15] R. Vesse, R.-M. Zettlemoyer, K. Ahmed, G. Moore, T. Pluskiewicz. (2013, May 6). dotNetR-DF An Open Source C# .Net Library for RDF [Online]. Available: http://www.dotnetrdf.org/ (downloaded 27 July 2013).