

## NRC Publications Archive Archives des publications du CNRC

### Weighted partonomy-taxonomy trees with local similarity measures for semantic buyer-seller match-making

Yang, Lu; Ball, Marcel; Bhavsar, Virendrakumar C.; Boley, Harold

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. /  
La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.

#### **Publisher's version / Version de l'éditeur:**

*Journal of Business and Technology, 1, 1, 2005-10*

#### **NRC Publications Archive Record / Notice des Archives des publications du CNRC :**

<https://nrc-publications.canada.ca/eng/view/object/?id=71ad364a-317c-4c77-860e-943c423393be>

<https://publications-cnrc.canada.ca/fra/voir/objet/?id=71ad364a-317c-4c77-860e-943c423393be>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at

<https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

**Questions?** Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

**Vous avez des questions?** Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

```

<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<mods xmlns="http://www.loc.gov/mods/v3" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" version="3.7" xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-7.xsd">
  <titleInfo xml:lang="en">
    <title>Weighted partonomy-taxonomy trees with local similarity measures for semantic buyer-seller match-making</title>
  </titleInfo>
  <name type="personal">
    <namePart type="family">Yang</namePart>
    <namePart type="given">Lu</namePart>
    <role>
      <roleTerm authority="marcrelator" type="code">aut</roleTerm>
    </role>
  </name>
  <name type="personal">
    <namePart type="family">Ball</namePart>
    <namePart type="given">Marcel</namePart>
    <role>
      <roleTerm authority="marcrelator" type="code">aut</roleTerm>
    </role>
  </name>
  <name type="personal">
    <namePart type="family">Bhavsar</namePart>
    <namePart type="given">Virendrakumar C.</namePart>
    <role>
      <roleTerm authority="marcrelator" type="code">aut</roleTerm>
    </role>
  </name>
  <name authority="nrcpid" type="personal">
    <namePart type="family">Boley</namePart>
    <namePart type="given">Harold</namePart>
    <role>
      <roleTerm authority="marcrelator" type="code">aut</roleTerm>
    </role>
    <nameIdentifier type="nrcpid">urn:nrcpid:15101</nameIdentifier>
    <affiliation xml:lang="en">IT</affiliation>
    <affiliation xml:lang="fr">ITI</affiliation>
  </name>
  <typeOfResource>text</typeOfResource>
  <genre authority="marcgt">article</genre>
  <language>
    <languageTerm authority="iso639-2b" type="code">eng</languageTerm>
  </language>
  <originInfo>
    <dateIssued encoding="w3cdtf">2005-10</dateIssued>
    <publisher xml:lang="en">Atlantic Academic Press</publisher>
  </originInfo>
  <abstract xml:lang="en">A semantically enhanced weighted tree similarity algorithm for buyer-seller match-making is presented. First, our earlier global (structural) similarity measure over (product) partonomy trees is enriched by taxonomic semantics: Inner nodes can be labeled by classes whose partial subsumption order is represented as a background taxonomy tree that is used for class similarity computation. In particular, the similarity of any two classes can be defined

```

via the weighted length of the shortest path connecting them in that taxonomy. To enable similarity comparisons between specialized versions of the background taxonomy, we encode subtaxonomy trees into partonomy trees in a way that allows the direct reuse of our partonomy similarity algorithm and permits weighted (or 'fuzzy') taxonomic subsumption with no added effort. Second, leaf nodes can be typed and each type be associated with a local, special-purpose similarity measure realizing the semantics to be invoked when computing the similarity of any two of its instances. We illustrate local similarity measures with e-Business types such as "Currency", "Address", "Date", and "Price". For example, the similarity measure on "Date"-typed leaf node labels translates various notations for date instances into a normal form from which it linearly maps any two to their similarity value. Finally, previous adjustment functions, which prevent similarity degradation for our arbitrarily wide and deep trees, are enhanced by smoother functions that evenly compensate intermediate similarity values.</abstract>

<abstract xml:lang="fr">Nous présentons ici un algorithme de similarité à arbre pondéré amélioré sémantiquement en vue de l'appariement des acheteurs et des vendeurs. Tout d'abord, nous enrichissons notre mesure antérieure globale (structurale) de la similarité sur des arbres de partonomie (produits) grâce à la sémantique taxonomique : les nœuds internes peuvent être étiquetés au moyen de classes dont l'ordre de subsumption partiel est représenté sous la forme d'un arbre de taxonomie de référence qui est utilisé pour le calcul de la similarité entre les classes. En particulier, on peut définir la similarité entre deux classes quelconques au moyen de la longueur pondérée du chemin le plus court qui les relie dans cette taxonomie. Afin de permettre d'établir des comparaisons de similarité entre des versions spécialisées de la taxonomie de référence, nous codons les arbres de sous taxonomie en arbres de partonomie de façon à permettre la réutilisation directe de notre algorithme de similarité partonomique et à autoriser une subsumption taxonomique pondérée (ou « floue »), sans travail additionnel. En second lieu, les nœuds feuilles peuvent être typés, et chaque type peut être associé à une mesure locale spécialisée de la similarité qui implémente la sémantique qui doit être appelée lors du calcul de la similarité de deux de ses instances quelconques. Nous illustrons la mesure de la similarité locale avec des types tirés du commerce électronique comme « Monnaie », « Adresse », « Date » et « Prix ». Ainsi, la mesure de la similarité des étiquettes des nœuds feuilles du type « Date » se traduit en diverses notations pour l'instance de la date dans une forme normale à partir de laquelle deux instances quelconques sont affectées linéairement à leur valeur de similarité. Enfin, les fonctions d'ajustement antérieures, qui empêchent la dégradation de la similarité dans un arbre d'une largeur et d'une profondeur arbitraires, sont améliorées par des fonctions plus régulières qui compensent de façon uniforme les valeurs de similarité intermédiaires.</abstract>

<note type="nrc pub">yes</note>

<note displayLabel="system details" type="admin" xml:lang="en">machine converted author identifier PE to PID, February 2012</note>

<relatedItem type="host">

<titleInfo xml:lang="en">

<title>Journal of Business and Technology</title>

</titleInfo>

<originInfo>

<dateIssued encoding="w3cdtf">2005-10</dateIssued>

<publisher xml:lang="en">Atlantic Academic Press</publisher>

</originInfo>

<part>

<detail type="volume">

<number>1</number>

</detail>

<detail type="issue">

<number>1</number>

</detail>

</part>

```
</relatedItem>
<identifier displayLabel="nrcc_number" type="nrc_id">NRCC 48518</identifier>
<extension>
  <organization authority="nrcna" valueURI="urn:nrcna:IN26">
    <term xml:lang="en">NRC Institute for Information Technology</term>
    <term xml:lang="fr">Institut de technologie de l'information du CNRC</term>
  </organization>
  <organization authority="nrcna" valueURI="urn:nrcna:IN13">
    <term xml:lang="en">National Research Council of Canada</term>
    <term xml:lang="fr">Conseil national de recherches du Canada</term>
  </organization>
  <taxonomy authority="nrcna" valueURI="urn:nrcna:GX520">
    <term xml:lang="en">National Research Council of Canada</term>
    <term xml:lang="fr">Conseil national de recherches du Canada</term>
  </taxonomy>
  <taxonomy authority="nrcna" valueURI="urn:nrcna:GX1041">
    <term xml:lang="en">NRC Publications Archive (NPArc)</term>
    <term xml:lang="fr">Archives des publications du CNRC (NPArc)</term>
  </taxonomy>
  <program authority="nrcna" valueURI="urn:nrcna:PG215">
    <term xml:lang="en">Intelligent Internet Applications</term>
    <term xml:lang="fr">Applications intelligentes pour l'internet</term>
  </program>
</extension>
<identifier type="nparc_id">8913970</identifier>
<recordInfo>
  <recordIdentifier source="tdr2">71ad364a-317c-4c77-860e-943c423393be</recordIdentifier>
</recordInfo>
</mods>
```