Bottleneck Structure Graphs in Multidomain Networks: Introduction and Requirements for ALTO

Jordi Ros-Giralt, Sruthi Yellamraju, Qin Wu, Richard Yang, Luis Contreras, Kai Gao, Jensen Zhang

I-Draft: draft-giraltyellamraju-alto-bsg-multidomain

https://datatracker.ietf.org/doc/draft-giraltyellamraju-alto-bsg-multidomain/

IETF Plenary 115
ALTO WG Session
11/11/2022
Table of Contents

- Brief Introduction to Bottleneck Structures
- Computing bottleneck structures under partial information
- Requirements discussion
Brief Introduction to Bottleneck Structures
Framework and Implementation Details in the Following I-Drafts and Papers


See also the IETF 115 PANRG sessions from yesterday (11/10/2022) for complete details:

10:05 20m  Computing Path Metrics Using Bottleneck Structure Graphs, draft-giraltyellamraju-alto-bsg-requirements  Jordi Ros-Giralt, Qin Wu
10:25 20m  Computing Bottleneck Structure Graphs in Multi-Domain Networks, draft-giraltyellamraju-alto-bsg-multidomain  Jordi Ros-Giralt, Qin Wu
Computing Bottleneck Structures Under Partial Information
Bottleneck Structures Under Partial Information: Problem Statement

Network with a single AS:
Bottleneck Structures Under Partial Information: Problem Statement

Network with a single AS:

Bottleneck Structure (Path Gradient Graph):
Bottleneck Structures Under Partial Information: Problem Statement

Network with a single AS:

Network with multiple ASs:

Bottleneck Structure (Path Gradient Graph):
Bottleneck Structures Under Partial Information: Problem Statement

Network with a single AS:

Bottleneck Structure (Path Gradient Graph):

Network with multiple ASs:

Incorrect substructure
Bottleneck Structures Under Partial Information: Problem Statement

Network with a single AS:

Network with multiple ASs:

Bottleneck Structure (Path Gradient Graph):

Bottleneck Structure of AS2:

Incorrect substructure

Correct substructure (got lucky)
Proposed Distributed Border Protocol

- **Convergence** (key idea): Sharing one metric per path is enough to ensure convergence to the correct bottleneck substructures.

- **Scalability**: Focuses on building the path gradient graph (see draft-giraltyellamraju-alto-bsg-requirements), requires only per-path state (not per-flow).

- **Privacy**: Does not require sharing sensitive flow information. Only one metric (scalar) per path is shared with the neighbor ASs.
Description of the Distributed Protocol

The algorithm run by each autonomous system AS$_i$, $1 \leq i \leq |A|$, consists of two independently executed events as follows:

**Event: TIMER**

- Every $s$ seconds, perform the following tasks:
  1. $B(A_i) = \text{COMPUTE_BOTTLENECK_SUBSTRUCTURE}(L(A_i), \text{PL}(A_i), C(A_i), \text{PM}(A_i))$;
  2. $\text{PM}(A_i)(p) = B(A_i).\text{BW}(p)$, for all $p$ in $P(A_i)$;
  3. For all $A_j$ in $N(A_i)$:
     3.1 Send to $A_j$ a $\text{PATH_METRIC_ANNOUNCEMENT}$ message including $(AS_i, \text{PM}(A_i))$;

**Event: PATH_METRIC_EXCHANGE**

- Upon receiving a $\text{PATH_METRIC_ANNOUNCEMENT}$ from AS$_j$ carrying $(AS_j, \text{PM}(A_j))$:
  1. $\text{PM}(A_i)(p) = \min\{\text{PM}(A_i)(p), \text{PM}(A_j)(p)\}$, for all $p$ in $P(A_i)$ and $p$ in $P(A_j)$;
Description of the Distributed Protocol

Procedure: \text{COMPUTE\_BOTTLENECK\_SUBSTRUCTURE}(L, PL, C, PM):

1. \text{i} = 0; L_0 = L; PL_0 = PL;

2. While True:
   2.1. \text{B}_i = \text{COMPUTE\_BOTTLENECK\_STRUCTURE}(\text{L}_i, \text{PL}_i, \text{C});
   2.2. If \text{B}_i.BW(p) == \text{PM}(p) for all path p in \text{PL}_i:
       2.2.1. Break;
   2.3. For all path p in \text{PL}_i such that \text{B}_i.BW(p) > \text{PM}(p):
       2.3.1. If \text{PL}_i[p] has no virtual link:
       2.3.1.1. Add a new virtual link v to the set of links \text{PL}_i[p];
       2.3.1.2. Add virtual link v to the set \text{L}_i;
       2.3.2. Set \text{C}(v) = \text{PM}(p);
   2.2. \text{i} = \text{i} + 1;
   2.5. \text{L}_i = \text{L}_{\text{i-1}};
   2.6. \text{PL}_i = \text{PL}_{\text{i-1}};

3. Return \text{B}_i;
Termination and Convergence Conditions

**Termination Condition.** When the output of an autonomous system’s local computation of the bottleneck structure is in agreement with the Path Metric Dictionary it maintains in coordination with its neighbors, the algorithm terminates:

\[ B_i.BW(p) = PM(p) \text{ for all path } p \text{ in } PL_i \]

**Convergence Condition.** Upon termination, the Path Metric Dictionaries of all the autonomous systems are in agreement with each other:

\[ PM(A_i)(p) = PM(A_j)(p) \text{ for all } p \text{ in } A_i, p \text{ in } A_j, A_i \text{ in } A \text{ and } A_j \text{ in } A \]
Example: Global Convergence to the Correct Bottleneck Substructures
Example: Global Convergence to the Correct Bottleneck Substructures
Example: Global Convergence to the Correct Bottleneck Substructures
Example: Global Convergence to the Correct Bottleneck Substructures

Convergence Condition

\[
\text{PM}(A_i)(p) = \text{PM}(A_j)(p)
\]
for all \( p \in A_i \), \( p \in A_j \), \( A_i \in A \) and \( A_j \in A \)

<table>
<thead>
<tr>
<th>p</th>
<th>PM(A1)(p)</th>
<th>PM(A2)(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1</td>
<td>8.3</td>
<td>--</td>
</tr>
<tr>
<td>p2</td>
<td>16.6</td>
<td>--</td>
</tr>
<tr>
<td>p3</td>
<td>8.3</td>
<td>--</td>
</tr>
<tr>
<td>p4</td>
<td>16.6</td>
<td>16.6</td>
</tr>
<tr>
<td>p5</td>
<td>--</td>
<td>75</td>
</tr>
<tr>
<td>p6</td>
<td>8.3</td>
<td>8.3</td>
</tr>
</tbody>
</table>

Sharing of PMDs using a PATH_METRIC_ANNOUNCEMENT message between neighbour ASs

Global Bottleneck Structure:
Example: Global Convergence to the Correct Bottleneck Substructures
Requirements discussion
Requirements Discussion

To implement the proposed distributed protocol using ALTO, two broad requirements are necessary:

- **Requirement 1**: The capability for each ALTO server to compute bottleneck substructures of its own AS.
- **Requirement 2**: The capability for each ALTO server to communicate with its neighboring ASs.

### 4.1. Requirement 1: Computation of Bottleneck Substructures

The requirements for an ALTO server to compute the bottleneck substructure of its associated AS are the same as the requirements to compute the bottleneck structure in the case the network consists of a single autonomous system. These requirements are discussed in the Requirements Section of [I-D.draft-giraltyellamraju-alto-bsg-requirements]. Refer to this document for further details.

### 4.2. Requirement 2: Communication Between Neighboring ASs

The **TIMER** event executed by each ALTO server needs to periodically transmit a **PATH_METRIC_ANNOUNCEMENT** message to its neighboring ASs. This leads to the following requirement:

- **Requirement 2.1**: ALTO servers managing neighboring ASs need to be reachable to each other.
- **Requirement 2.2**: The sharing of algorithmic state between ALTO servers requires extending the base ALTO protocol to support server-to-server communication semantics.
Discussion Q&A

Thank you