Link-State Routing Optimization for Compound Autonomous Systems in the Internet

Juan Antonio Cordero

École Polytechnique, France
Laboratoire d’Informatique (LIX)
Équipe HIPERCOM
Towards a more flexible Internet

ARPANET, 1969
(from http://som.csudh.edu/cis/lpress/history/arpamaps/)

Internet connections in the US, 2009
(map by Z. Deretsky, NSF, adapted from maps by C. Harrison, Human-Computer Interaction Institute of CMU, www.chrisharrison.net)
Towards a more flexible Internet

- **Size**: connections, hosts, users…
  
  ~ 850 Mhosts
  (Internet Domain Survey Count, July 2011)

- **Routing architecture**: ARPANET → Internet

- **Complexity**: types of interconnected networks
Towards a more flexible Internet

- Autonomous Systems

- Routing autonomy
- Inside / outside

Host
Router in an AS
Router between ASes
Towards a more flexible Internet

- Wireless, mobile, ad hoc

- 1985: US FCC allows unlicensed use of wireless (ISM) spectrum

- 1990: IEEE launches 802.11 (Wi-Fi) standardization group

- 1997: IETF defines Mobile Ad hoc Networking (MANET)
Towards a more flexible Internet

- Wireless ad hoc networks
  - Topology unknown \textit{a priori}
  - May change \textit{unpredictably} during network operation
Towards a more flexible Internet

- **Applicability** of wireless and ad hoc networks

*Rescue and recovery scenarios*
Towards a more flexible Internet

- **Applicability** of wireless and ad hoc networks

*Vehicular networks (VANETs)*

*Rescue and recovery scenarios*
Towards a more flexible Internet

- **Applicability** of wireless and ad hoc networks

- **Wireless sensor networks (WSNs)**

- **Vehicular networks (VANETs)**

- **Rescue and recovery scenarios**
Towards a more flexible Internet

- **Applicability** of wireless and ad hoc networks

  - Wireless sensor networks (WSNs)
  - Rescue and recovery scenarios

  - Vehicular networks (VANETs)
  - Spontaneous / community mesh networking
Towards a more flexible Internet

- **Issues** of wireless ad hoc networking
  - Wireless "links", partly shared medium
  - Unreliability in wireless communication
  - Topology dynamism and router mobility

Degradation of communication vs wireless links
Routing in Compound Autonomous Systems

- Compound Autonomous System
  - Fixed networks + Wireless ad hoc networks

Rest of the Internet
Routing in Compound Autonomous Systems

- Compound Autonomous System
  - Fixed networks + Wireless ad hoc networks

Rest of the Internet

Compound Autonomous System
Routing in Compound Autonomous Systems

Key questions

- Why addressing routing in compound Autonomous Systems?
- One or several routing solutions for a compound AS?
Routing in Compound Autonomous Systems

- Key questions
  - Why addressing routing in compound Autonomous Systems?
    - Currently deployed IGPs do not work in wireless ad hoc networks
  - One or several routing solutions for a compound AS?
Routing in Compound Autonomous Systems

Key questions

- Why addressing routing in compound Autonomous Systems?
- One or several routing solutions for a compound AS?
Routing in Compound Autonomous Systems

- Extension of an already used IGP for wireless ad hoc operation
- Open Shortest Path First (OSPF) : one of the main link-state IGPs
Link-State Routing Optimization for Compound Autonomous Systems in the Internet

**Link-State Routing**

- Link-state Database (LSDB)
- Shortest Path Tree (SPT)
- Routing Table

(Dijkstra)  (next-hop)
Link-State Routing Optimization for Compound Autonomous Systems in the Internet

Link-State Routing

Link-State Advertisements (LSAs)

Link-state Database (LSDB)  \rightarrow  Shortest Path Tree (SPT)  \rightarrow  Routing Table

(Dijkstra)  \quad (\text{next-hop})
Link-State Routing Optimization for Compound Autonomous Systems in the Internet

Link-State Routing

**LSA Acquisition**

- Routers advertise *all* their neighbors in their own LSA
- Routers retransmit *every* LSA they receive *immediately*
- Routers advertise all their neighbors in their own LSA
- Routers retransmit every LSA they receive immediately
- Every pair of routers have the same information in their LSDBs
- All routers receive LSAs from every other router in the network
Link-State Routing

**Link State Operations**

- Routers advertise *all* their neighbors in their own LSA
- Routers retransmit *every* LSA they receive *immediately*
- Every pair of routers have *the same* information in their LSDBs
- All routers receive LSAs from *every other* router in the network

**Topology Description**

**Flooding**

**LSDB Synchronization**
### Link-State Routing

#### Link State Routing over Wireless Ad hoc Networks

<table>
<thead>
<tr>
<th><strong>Objective</strong></th>
<th><strong>Open issues</strong></th>
<th><strong>Reduction goals</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Topology view</td>
<td>Routes towards destinations</td>
<td>Shortest paths</td>
</tr>
<tr>
<td>Flooding</td>
<td>Dominating set of forwarders</td>
<td>Wireless collisions</td>
</tr>
<tr>
<td>Synchronization</td>
<td>Synchronized paths</td>
<td>LSDB exchange</td>
</tr>
</tbody>
</table>

### Open issues

- Reduction goals
- Adv. links, updates
- Txs, latency
- Synchronizations

### References


### Objectives

- **Topology view**: Routes towards destinations
- **Flooding**: Dominating set of forwarders
- **Synchronization**: Synchronized paths

### Open Issues

- Shortest paths
- Wireless collisions
- LSDB exchange

### Reduction Goals

- Adv. links, updates
- Txs, latency
- Synchronizations

---

Separate optimization of each link-state operation over MANETs

---


Contributions

*Optimization techniques for link-state routing over MANETs*

- Synchronized Link Overlay – Triangular (SLOT)
- Multi-Point Relays (MPR)
- Smart Peering (SP)

*Application to OSPF*

- Implementation/Simulation of Extensions for MANETs
- Experiments in a Compound Internetwork
Contributions

Optimization techniques for link-state routing over MANETs

- Synchronized Link Overlay – Triangular (SLOT)
  - Multi-Point Relays (MPR)
  - Smart Peering (SP)

Application to OSPF

- Implementation/Simulation of Extensions for MANETs
- Experiments in a Compound Internetwork
Optimization techniques

Synchronized Link Overlay

- Based on the Relative Neighborhood Graph (RNG)

Given a set of points $S \subset \mathbb{R}^n$, $x, y \in S$,

$$(x, y) \in RNG(S) \iff \forall z \in S, d(x, z), d(z, y) > d(x, y)$$
Optimization techniques

**Synchronized Link Overlay**

- Two **variants** considered, depending on the **metric** in use:
  - SLOT-D ($\mathbb{R}^n$ distance)
  - SLOT-U (hop count)

---


Link-State Routing Optimization for Compound Autonomous Systems in the Internet

Optimization techniques

**Synchronized Link Overlay**

Example

- Network graph
Optimization techniques

Synchronized Link Overlay

Example

– SLOT-U subgraph
Optimization techniques

**Synchronized Link Overlay**

- SLOT-D subgraph
Optimization techniques

Synchronized Link Overlay

Average number of links per node

\[ M_{\text{SLOT-}U}(\nu) \leq 3.6039 \]
\[ M_{\text{SLOT-}D}(\nu) \leq 2.5575 \]

- Connected overlays
- Bounded overlay size

![Graph showing average number of links per node with lines for SLOT-U and SLOT-D, and markers for 3.60 and 2.56]

Density (nodes/\text{u2})
Average link creation rate for a node speed $s$

$$s = 5 \frac{u}{\text{sec}}$$

- Improves overlay stability

$$V_{SLOT-U}(s, v) \approx O(s)$$

$$V_{SLOT-D}(s, v) \approx O\left(s \sqrt{v}\right)$$
Contributions

*Optimization techniques for link-state routing over MANETs*

- Synchronized Link Overlay – Triangular (SLOT)
- **Multi-Point Relays (MPR)**
- Smart Peering (SP)

*Application to OSPF*

- Implementation/Simulation of Extensions for MANETs
- Experiments in a Compound Internetwork
 Optimization techniques

Multi-Point Relays

- Definition

- 2-hop neighborhood
- Designed for flooding
- Adaptation for topology selection


Optimization techniques

Multi-Point Relays

- Definition

- 2-hop neighborhood

- Designed for flooding

- Adaptation for topology selection

(Enhanced) Path MPR


Optimization techniques

Multi-Point Relays

- Not adapted for LSDB synchronization
- High link change rate

![Graph showing the average lifetime of multipoint relays and bidirectional neighbors for a fixed size grid at 5 m/s](image)

(GTNetS simulations, mobile network)

Optimization techniques

Multi-Point Relays

- Not adapted for LSDB synchronization
- High link change rate

![Graph]

Average lifetime of multipoint relays and bidirectional neighbor (Fixed size grid, 5 m/s)

- Multi-point relays
- Bidirectional neighbors

(GTNetS simulations, mobile network)

Contributions

Optimization techniques for link-state routing over MANETs

- Synchronized Link Overlay – Triangular (SLOT)
- Multi-Point Relays (MPR)
- Smart Peering (SP)

Application to OSPF

- Implementation/Simulation of Extensions for MANETs
- Experiments in a Compound Internetwork
Optimization techniques

Smart Peering

Definition

- Depends on the network dynamics
- Priority to synchronization with routers maintaining stable links

### Optimization techniques

- **Summary**

<table>
<thead>
<tr>
<th></th>
<th>SLOT</th>
<th>MPR</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Requirements</strong></td>
<td>1-hop neighbors</td>
<td>2-hop neighbors</td>
<td>Global topology knowledge</td>
</tr>
<tr>
<td><strong>Topology Selection</strong></td>
<td></td>
<td>Enhanced Path MPR</td>
<td></td>
</tr>
<tr>
<td><strong>Flooding</strong></td>
<td></td>
<td>Full coverage in 2 hops</td>
<td></td>
</tr>
<tr>
<td><strong>LSDB Synchronization</strong></td>
<td>Density bounded Reduced change rate (not dependent on density for SLOT-U)</td>
<td>High link change rate ($\Rightarrow$ persistency)</td>
<td>Sensitive to mobility Priority to stable links</td>
</tr>
</tbody>
</table>
Contributions

*Optimization techniques for link-state routing over MANETs*

- Synchronized Link Overlay – Triangular (SLOT)
- Multi-Point Relays (MPR)
- Smart Peering (SP)

*Application to OSPF*

- Implementation/Simulation of Extensions for MANETs
- Experiments in a Compound Internetwork
Application to OSPF

Open Shortest Path First

- Two routing principles

  - User data is forwarded over shortest paths
  
  - User data & control traffic is sent over synchronized links

… suitable over MANETs?
Application to OSPF

OSPF over MANET

- IETF OSPF MANET extensions
  - Multipoint Relays (MPR-OSPF)   RFC 5449
  - MPRs, Smart Peering (OR / SP)  RFC 5820

- Additional extensions
  - MPR + SP
  - SLOT-OSPF
  - Persistent Variants of MPR-OSPF


### Application to OSPF

**OSPF over MANET**

- **Topology Selection**
  - MPR-OSPF: Path MPR
  - OR/SP: SP
  - MPR+SP: Path MPR (+ SP)
  - SLOT-OSPF: Path MPR

- **Flooding**
  - MPR-OSPF: Flooding MPR
  - OR/SP: MPR ○ SP
  - MPR+SP: Flooding MPR
  - SLOT-OSPF: Flooding MPR

- **LSDB Synchr.**
  - MPR-OSPF: MPR
  - OR/SP: SP
  - MPR+SP: SP
  - SLOT-OSPF: SLOT

---

<table>
<thead>
<tr>
<th></th>
<th>MPR-OSPF (RFC5449)</th>
<th>OR/SP (RFC5820)</th>
<th>MPR+SP</th>
<th>SLOT-OSPF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Topology Selection</strong></td>
<td>Path MPR</td>
<td>SP</td>
<td>Path MPR (+ SP)</td>
<td>Path MPR</td>
</tr>
<tr>
<td><strong>Flooding</strong></td>
<td>Flooding MPR</td>
<td>MPR ○ SP</td>
<td>Flooding MPR</td>
<td>Flooding MPR</td>
</tr>
<tr>
<td><strong>LSDB Synchr.</strong></td>
<td>MPR</td>
<td>SP</td>
<td>SP</td>
<td>SLOT</td>
</tr>
</tbody>
</table>
It is beneficial to preserve \textit{shortest paths} for user data in MANETs.
Synchronizing (all) shortest paths is costly in MANETs…

All links included in the SPT are declared adjacent
Application to OSPF

OSPF over MANET

Only Smart Peering links are adjacent (MPR+SP)

All links included in the SPT are adjacent (MPR-OSPF)

… and does not improve significantly the quality of routing

Application to OSPF

OSPF over MANET

Delivery ratio
(Fixed size grid, 5 m/s)

MPR+SP vs. SLOT-OSPF

- Advertise synchronized links?

MPR+SP vs. SLOT-OSPF
Contributions

**Optimization techniques for link-state routing over MANETs**

- Synchronized Link Overlay – Triangular (SLOT)
- Multi-Point Relays (MPR)
- Smart Peering (SP)

**Application to OSPF**

- Implementation/Simulation of Extensions for MANETs

- **Experiments in a Compound Internetwork**
Testbed: compound internetwork, 6 computers

- 3 with wired interfaces
- 3 with wireless interfaces
- 2 with wired and wireless interfaces

Routing based on OSPF

- OSPF for IPv6 (RFC 5340)
- MPR-OSPF (RFC 5449)
Application to OSPF

OSPF in Wired/Wireless Internetworks

Distribution of computers at LIX
Application to OSPF

OSPF in Wired/Wireless Internetworks

Logical topologies considered in the internetwork
Application to OSPF

OSPF in Wired/Wireless Internetworks

Logical topologies considered in the internetwork

1 wired hop + 1 wireless hop

1 + 2

1 + 3
Application to OSPF

OSPF in Wired/Wireless Internetworks

- Quality of communication degrades linearly as packets traverse more wireless hops

Packet Delivery Ratio (PDR) vs Number of wireless hops
Application to OSPF

OSPF in Wired/Wireless Internetworks

Synchronization traffic is present during the whole lifetime of a wireless link.
Main contributions

- **Enable the Internet to exploit wireless ad hoc networking capabilities**
- **Framework** for analysis of link-state routing in compound ASes
- **Optimization** of link-state operations over MANETs
  - Proposal of new techniques
  - Theoretical analysis of existing techniques
  - Improvement / generalization of techniques
- **Implementation and evaluation in OSPF**
  - Development and simulations of OSPF MANET extensions
  - Experiments over OSPF in compound internetworks
Towards a wireless, more mobile Internet

Optimizations in other link-state routing protocols

Metrics in ad hoc and compound networks

- Beyond hop count: link reliability, distance, available bandwidth, maximum throughput (ETX)...
- Fixed / ad hoc coexistence requires a clear distinction between wired and wireless links

... beyond the principle of deterministic shortest path?
Questions ?
Backup slides
Autonomous Systems

- **RFC 975** (1986): “…a set of gateways, each of which can reach any other gateway in the same system using paths via gateways only in that system. The gateways of a system cooperatively maintain a routing data base using an interior gateway protocol…”

- **RFC 1930** (1996): “…a connected group of one or more IP prefixes [internetworks] (..) which has a SINGLE and CLEARLY DEFINED routing policy.”

- **RFC 1812** (1995): “…routers [inside an AS] may use one or more interior routing protocols…”
Definition (def. 1.5)

Let $a$, $b$ be network interfaces.

There is a link $l$ between $a$ and $b$, denoted by $l: a \rightarrow b$, iff $a$ is able to transmit data to $b$ and $b$ is able to receive such data, without the intervention of any other network interface.

Equivalence (def. A.1)

Let $l_1: s_1 \rightarrow d_1$, $l_2: s_2 \rightarrow d_2$ be links.

$l_1$ and $l_2$ are equivalent, and denoted as $l_1 \equiv l_2$ iff any of the following conditions is satisfied:

(i) $s_1 = s_2$, $d_1 = d_2$
(ii) $s_1 = s_2$, $d_1 \neq d_2$ and any pkt sent from $s_1$ to $d_1$ (via $l_1$) is also received by $d_2$ (via $l_2$), and viceversa
(iii) $s_1 \neq s_2$ and $\exists l_{12}^*: s_1 \rightarrow s_2$, $l_{21}^*: s_2 \rightarrow s_1$ | any pkt sent from $s_1$ to $d_1$ (via $l_1$) is also received by $s_2$ (via $l_{12}^*$) and $d_2$ (via $l_2$).

(Prop. A.1) Relation $\equiv$ is an equivalence relation.
Simulation Parameters

General Simulation Parameters

- 20 samples/experiment

- **Data traffic pattern**
  - Constant Bit Rate UDP flow
  - Packet size: 1472 bytes
  - Packet rate: 85 pkts/sec

- **Scenario**
  - Square grid
  - Grid size: 400x400 m

- **Node configuration**
  - Radio range: 150 m
  - Propagation: Two-ray
  - Wireless $\alpha$: 0.5
  - MAC protocol: IEEE 802.11b

- **Node mobility**
  - Random waypoint model
  - Pause: 0 sec
  - Speed: 0,5 m/s (ct.)

Performed Experiments

- Fixed size grid

**OSPF Configuration**

- **Standard Parameters**
  - HelloInterval: 2 sec
  - DeadInterval: 6 sec
  - RxsInterval: 5 sec
  - MinLSInterval: 5 sec
  - MinLSArrival: 1 sec
  - LSRefreshInterval: 20 sec

- **RFC 5449-like**
  - AckInterval: 1.8 sec

- **RFC 5820**
  - AckInterval: 1.8 sec
  - PushbackInterval: 2 sec
Networking Interfaces
- Wired: Digital Equipment Corp. DECchip 21140
- Wireless: Broadcom BCM4306 WLAN

Software
- OS: Ubuntu v.10.04, kernel 2.6.32
- Routing implementation: ospf6d daemon of Quagga/Zebra suite, v.0.99.15
- OSPF interface types
  - Wired: Point-to-point
  - Wireless: MANET, RFC 5449

OSPF Configuration
- HelloInterval: 2 sec
- DeadInterval: 10 sec
- RxmtInterval: 5 sec
- AckInterval: 2 sec
- Jitter (max.): 100 msec
- MinLSInterval: 5 sec
- MinLSArrival: 1 sec
- LSRefreshInterval: 60 sec

UDP Flows
- Nom. sender bitrate: 100 pkt/s
- Packet payload: 1024 B
- CBR real traffic rate: 300 kbps
- Flow duration: 5 min/flow

Measures
- Router starting: [0, 2] sec
- PDR of UDP flows: 60 samples
- Control traffic: 84 samples
Wireless $\alpha$

![Graph showing the probability of success over link length for different values of $\alpha$.]
MPR Heuristics

- **Input**: \( x, N(x), N_2(x) \)
- **MPR** = \( \emptyset \)
- **MPR** \( \leftarrow \{ n \in N(x) : \exists m \in N_2(x), m \text{ only covered by } n \} \)
- while (\( \exists \) uncovered 2-hop neighbors)
  MPR \( \leftarrow n \in N(x) : \text{covers max. # of uncovered 2-hop neighbors} \)
- **Output**: \( \text{MPR}(x, N(x), N_2(x)) \)
Enhanced Path MPR

\[ N'(x) = \{ n \in N(x) : m(x, n) = \text{dist}_2(x, n) \} \subseteq N(x) \]

\[ N'_2(x) = \{ n \in N(x) \cup N_2(x) \mid n \notin N'(x), \exists m \in N'(x) : m(n, m) + m(m, x) = \text{dist}_2(n, x) \} \]

\[ \subseteq N(x) \cup N_2(x) \]

\[ (E^2_x)' = \{ nm \in E(G) : n \in N'(x), m \in N'_2(x), m(x, n) + m(n, m) = \text{dist}_2(x, m) \} \]

\[ \cup \{ xn \in E(G) : n \in N'(x) \} \]
Persistency in MPR-OSPF (1)

<table>
<thead>
<tr>
<th></th>
<th>Adjacencies</th>
<th>Flooding</th>
<th>Topology</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPM</td>
<td>Persistent</td>
<td>Persistent</td>
<td>Path MPRs</td>
</tr>
<tr>
<td>PMP</td>
<td>Persistent</td>
<td>MPR Selectors</td>
<td>Persistent</td>
</tr>
<tr>
<td>PMM</td>
<td>Persistent</td>
<td>MPR Selectors</td>
<td>Path MPRs</td>
</tr>
<tr>
<td>MMM</td>
<td>MPRs</td>
<td>MPR Selectors</td>
<td>Path MPRs</td>
</tr>
</tbody>
</table>

(RFC 5449) (Non-pers.)
- The non-persistent configuration \((MMM)\) performs **significantly worse** than the other (partially persistent) configurations.

- The delivery achieved by current standard *RFC 5449* can be **improved** by implementing persistency also in flooding \((PPM)\) or in topology selection \((PMP)\).
Persistent adjacencies are **far more stable**, but persistency **increases the size** of the adjacent links set.
Overall, the cost (in overhead) of implementing persistency in flooding (PPM) is more significant than implementing it for topology selection (PMP)

The benefits of flooding persistency are roughly equivalent as those of topology selection persistency.

Non-persistent configuration (MMM) generates more overhead than other persistent confs. (PMM and PMP) due to adjacency unstability.
Flooding: Stating the problem

Wireless collisions (1)
The Jitter mechanism

- Flooding with jitter | Intuition

Wireless collisions (2)
Flooding with jitter: RFC 5148

- Received pkt at t=t₀
  - Assigns a jitter value j to all msgs of the pkt
  - N=1
  - Extracts N-th msg from the pkt
  - N-th msg needs to be forwarded?
    - Yes
      - Schedule tx at t=t₀+j
    - No
      - ⊤ Next N?
        - No
        - Yes
          - Send all msgs scheduled and not sent at t=t₂
  -调度 tx at t=t₁
  - t₂=t₁

Wireless collisions (3)
**The Jitter mechanism**

- **Objectives**
  - Theoretical analysis of the *impact* of jitter in flooding
    - (from the perspective of a *wireless interface*)

- **Focusing on:**
  - *Delay* introduced in forwarding
  - Changes in (out-) *packet transmission rate*
Given a jitter value $T$, what is the average delay before retransmission?

Two (extreme) cases:

- Retx of a packet only depends on further arrivals
- Retx of a packet depends (also) on all possible previous arrivals
Introduced delay (in average)