Geodiverse Multipath Communication
with Structural Multilevel Diversity
for the Resilient Future Internet

James P.G. Sterbenz*†
Джеймс Ф.Г. Штербэнз
Джеймс Стербенз

Dongsheng Zhang 张东升,
Yufei Cheng 成宇飞,
Anh Nguyên, Mohammed Alenazi

*Department of Electrical Engineering & Computer Science
Information Technology & Telecommunications Research Center
The University of Kansas
†School of Computing and Communications, Infolab 21
Lancaster University

jpgs@ittc.ku.edu
http://www.ittc.ku.edu/~jpgs
http://wiki.ittc.ku.edu/resilinets
Where is Kansas?

Geography Lesson

KANSAS

San Francisco

San Diego

Phoenix

Salt Lake City

Denver

Kansas City

OZ

Lawrence

OZ

KU

UK

Zürich

Oslo

Россия

서울

東京

香港

台北市

北京

台北市

東京

香港

北京

서울

東京

香港

台北市

北京

서울

東京

香港

台北市

北京

서울

東京

香港

台北市

北京
Multilevel Structural Diversity

Outline

- ResiliNets review
- Multilevel interrealm resilience
  - resilience to attackers
  - resilience to large scale disasters
Resilience and Survivability
Motivation and Definition

• Increasing reliance on network infrastructure
  ⇒ Increasingly severe consequences of disruption
  ⇒ Increasing attractiveness as target from bad guys

• Need *resilience*
  – provide and maintain acceptable service
  – in the face of faults and challenges to normal operation

• Challenges
  – ...
  – large scale disasters (natural and human-caused)
  – malicious attacks from intelligent adversaries
ResiliNets Initiative

Goals

• Understand network structure and vulnerabilities
  – develop new models and tools for analysis

• Develop ways to increase network resilience
  – improving existing networks under cost constraints
  – increase cost to attackers
  – Future Internet design
  – validate by analysis, simulation, and experimentation

• Funded primarily by
  – US NSF FIND and GENI programs and open call (with Medhi)
  – US DoD
  – EU FP6 and FP7 FIRE programme (with David Hutchison)
ResiliNets Strategy

$D^2R^2 + DR$

• Two phase strategy for resilience
• Real time control loop: $D^2R^2$
  – defend
    • passive
    • active
  – detect
  – remediate
  – recover
• Background loop: DR
  – diagnose
  – refine

[Wiki 2005, ComNet 2010]
ResiliNets Principles
High Level Grouping

- Prerequisites: to understand and define resilience
- Tradeoffs: recognise and organise complexity
- Enablers: architecture and mechanisms for resilience
- Behaviour: require significant complexity to operate
Resilience Principles
Redundancy, Diversity, Heterogeneity

- Diversity
  - mechanism (wired & wireless), provider, geographic path
- Multipath transport
  - spreading (erasure code) or as hot-standby
Multilevel Structural Diversity
Multilevel Interrealm Resilience

• ResiliNets review
• Multilevel interrealm resilience
  – resilience to attackers
  – resilience to large scale disasters
Multilevel Network Topology
Example: Sprint L3 IP PoP Topology
Multilevel Network Topology

Example: Sprint L3 overlay on L2.5
Multilevel Network Topology

Example: Sprint L2.5 MPLS PoP Topology
Multilevel Network Topology
Example: Sprint L2.5 overlay on L2/1
Multilevel Network Topology
Example: Sprint L1 Physical Fiber Topology
Multilevel Network Topology

Example: Sprint L1–3 Topology
Complex Network Topology
KU-TopView Topology Viewer

L1 Sprint fiber visualisation ↔ adjacency matrices
• KU-CSM Challenge Simulation Module
  – challenge specification describes challenge scenario
  – network coordinates provide node geo-locations
  – adjacency matrix specifies link connectivity
  – input to conventional ns-3 simulation run
  – generates trace to plot results with KU-gpWrapper [RNDM 2010]
Challenge Simulation

Challenge Types

- Challenge types
  - node or link down
    - random or attack (deg, betweenness, ...)
  - area based challenge
    - \( n \)-sided polygon: \((x_0, y_0), \ldots, (x_{n-1}, y_{n-1})\)
    - circle centered at \((x_0, y_0)\) with radius \(r\)
  - wireless link attenuation or jamming
  - traffic attacks (DoS and DDoS)

- Challenge characteristics
  - type (e.g. wired/wireless)
  - class (e.g. important peering node)
  - dynamic: interval \((t_i, t_j)\), trajectory
Multilevel Structural Diversity

Resistance to Attackers

- ResiliNets review
- Multilevel interrealm resilience
  - resilience to attackers
  - resilience to large scale disasters
Multilevel Network Analysis
Abstraction of Internet Topology

[AS 1] [AS level topology] [AS 2]
[AS 3] [AS 4]

[ISP 1] [router level topology] [ISP 2]
[IXP 1] [IXP 2] [IXP 3] [IXP 4]

[ISP 3] [ISP 4]

[physical level topology]

[DRCN 2013]
Multilevel Network Analysis
Multilevel Graph Model

- Multilevel model for unweighted & undirected graphs
- Two requirements for multilevel graph model:
  - nodes at the above level are subset of lower level
  - nodes that are disconnected below are disconnected above

---

Connected network
Disconnected network
Partitioned network
### Resilience Analysis

#### Graph-Theoretic Properties of Networks

<table>
<thead>
<tr>
<th>Topology</th>
<th>Sprint Physical</th>
<th>Sprint Logical</th>
<th>AT&amp;T Physical</th>
<th>AT&amp;T Logical</th>
<th>US Highways</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nodes</td>
<td>263</td>
<td>28</td>
<td>361</td>
<td>107</td>
<td>400</td>
</tr>
<tr>
<td>Number of links</td>
<td>311</td>
<td>76</td>
<td>466</td>
<td>140</td>
<td>540</td>
</tr>
<tr>
<td>Maximum degree</td>
<td>6</td>
<td>14</td>
<td>7</td>
<td>23</td>
<td>7</td>
</tr>
<tr>
<td>Average degree</td>
<td>2.37</td>
<td>5.43</td>
<td>2.58</td>
<td>2.62</td>
<td>2.7</td>
</tr>
<tr>
<td>Degree assortativity</td>
<td>-0.17</td>
<td>-0.23</td>
<td>-0.16</td>
<td>-0.4</td>
<td>0.11</td>
</tr>
<tr>
<td>Node closeness</td>
<td>0.07</td>
<td>0.48</td>
<td>0.08</td>
<td>0.3</td>
<td>0.08</td>
</tr>
<tr>
<td>Clustering coefficient</td>
<td>0.03</td>
<td>0.41</td>
<td>0.05</td>
<td>0.09</td>
<td>0.05</td>
</tr>
<tr>
<td>Algebraic connectivity</td>
<td>0.0053</td>
<td>0.6844</td>
<td>0.0061</td>
<td>0.1324</td>
<td>0.0059</td>
</tr>
<tr>
<td>Network diameter</td>
<td>37</td>
<td>4</td>
<td>37</td>
<td>6</td>
<td>40</td>
</tr>
<tr>
<td>Network radius</td>
<td>19</td>
<td>2</td>
<td>19</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>Average hop count</td>
<td>14.78</td>
<td>2.19</td>
<td>13.57</td>
<td>3.38</td>
<td>13.34</td>
</tr>
<tr>
<td>Node betweenness</td>
<td>11159</td>
<td>100</td>
<td>15970</td>
<td>2168</td>
<td>22798</td>
</tr>
<tr>
<td>Link betweenness</td>
<td>9501</td>
<td>27</td>
<td>14270</td>
<td>661</td>
<td>18585</td>
</tr>
</tbody>
</table>
Multilevel Resilience
Effect of Physical Failures on L3 Topology

- Attacks against physical infrastructure
  - based on centrality (importance) metrics
  - adaptive recomputes metrics after each node failure
- Analysis of impact on higher layer flows
  - heuristics to add elements under cost constraints
Multilevel Structural Diversity
Resilience to Large-Scale Disasters

- ResiliNets review
- Challenge Taxonomy
- Multilevel interrealm resilience
  - resilience to attackers
  - resilience to large scale disasters
Simulation Analysis

Example: Multilevel Analysis of Disaster

- Hurricane disaster in New Orleans area
- Destruction of physical infrastructure
- Effect on IP-layer network services
Resilience Analysis
Path and Graph Diversity

- **Path diversity**
  - measure of links and nodes in common
- **EPD**: effective path diversity \([0,1]\)
  - normalised diversity with respect to a single shortest path
  - measure of E2E flow resilience
- **TGD**: total graph diversity is average of EPD
  - for all pairs: quantifies available diversity in graph

\[
D(P_k) = 1 - \frac{|P_k \cap P_0|}{|P_0|}
\]

[DRCN 2009]
Resilience Analysis
Path and Graph Diversity with Distance Metric

• cTGD: compensated TGD
  – weighted to be predictive of flow robustness [RNDM 2010]
  – algebraic connectivity also fair predictor or flow robustness

• GeoPath diversity
  – distance $d$ between paths beyond source and destination
  – GeoResLSR: $(k, d, [s,t])$ multipath geographic routing
    • number of paths $k$
### Resilience Analysis

**Compensated Total Graph Diversity**

<table>
<thead>
<tr>
<th>Metric Network</th>
<th>surv</th>
<th>deg</th>
<th>cTGD</th>
<th>TGD</th>
<th>clus coef</th>
<th>dia</th>
<th>hop cnt</th>
<th>clse</th>
<th>nod btw</th>
<th>link btw</th>
</tr>
</thead>
<tbody>
<tr>
<td>full mesh</td>
<td>01</td>
<td>01</td>
<td>01</td>
<td>01</td>
<td>01</td>
<td>01</td>
<td>01</td>
<td>01</td>
<td>01</td>
<td>01</td>
</tr>
<tr>
<td>Level3</td>
<td>02</td>
<td>02</td>
<td>02</td>
<td>02</td>
<td>02</td>
<td>04</td>
<td>02</td>
<td>03</td>
<td>10</td>
<td>09</td>
</tr>
<tr>
<td>AboveNet</td>
<td>03</td>
<td>03</td>
<td>03</td>
<td>08</td>
<td>03</td>
<td>03</td>
<td>03</td>
<td>02</td>
<td>05</td>
<td>03</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ring</td>
<td>15</td>
<td>07</td>
<td>15</td>
<td>13</td>
<td>15</td>
<td>09</td>
<td>15</td>
<td>15</td>
<td>04</td>
<td>08</td>
</tr>
<tr>
<td>AT&amp;T L1</td>
<td>16</td>
<td>07</td>
<td>16</td>
<td>03</td>
<td>13</td>
<td>10</td>
<td>16</td>
<td>16</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>Sprint L1</td>
<td>17</td>
<td>07</td>
<td>17</td>
<td>06</td>
<td>14</td>
<td>10</td>
<td>17</td>
<td>17</td>
<td>12</td>
<td>16</td>
</tr>
</tbody>
</table>

- cTGD much better predictor of flow robustness
  - cTGD with $\alpha = 0.25$ perfect predictor for these 17
- 13 real networks plus 4 regular topologies

[RNDM 2011]
ResiliNets Protocols
Cross-Layer Model: Generic

- **Knobs** $K_{i \rightarrow i-1} = \{k_i\}$ influence behaviour to levels below
- **Dials** $D_{i+1 \leftarrow i} = \{d_i\}$ expose characteristics to upper levels
- **Levels** (of significance to ResiliNets)
  - 8: social
  - 7: application
  - 4: end-to-end transport
  - 3i: inter-realm (domain)
  - 3r: routing
  - 3t: logical topology
  - 2: hop-by-hop links
  - 1: physical topology
Resilient Transport: ResTP
Overview

• ResTP: Resilient Transport Protocol
  – flexible and composable [ala TP++ [Feldmeier, McAuley]]

• Flexible and composable
  – flow setup and management
    • including multipath support
  – error control
  – transmission (flow and congestion) control

• Cross-layered
  – applications specify service and threat model
  – behaviour based on path characteristics
  – specifies path requirements to GeoDivRP
Resilient Transport: ResTP

Reliability Modes

- Reliability (combination of flow and error modes)
  - full reliability: E2E 3-way handshake and ACKs
  - nearly-reliable: custody transfer at GW with e2e ACKs
  - quasi-reliable: E2E FEC giving statistical reliability
  - none (flow): connection oriented best effort
  - none (datagram): connectionless best effort (UDP-like)

- Chosen using cross-layering
  - service specification and threat model from application
  - path characteristics from lower layers
Resilient Transport: ResTP

Flow Modes

- Multiple flow modes
  - hard connections (3-way handshake)
  - opportunistic connections (signalling overlaps data)
  - custody transfer at realm boundaries (for DTNs)
    - AeroTP subset of ResTP uses this
  - soft-state flows
  - signalled flow with datagrams
  - individual datagrams
Resilient Transport: ResTP

Error Control

- Multipath modes
  - alternate path added on-demand
  - alternate path as hot-standby
  - erasure coding across \( k \) paths (typically \( k=3 \))
    - best coding for large skew?

- Per subflow modes
  - ARQ for reliable service
    - SACK, MACK, NAK, SNACK (SCPS-style)
  - HARQ for reliable service on lossy path
  - adaptive FEC for quasireliable service
  - none for unreliable service
Resilient Transport: ResTP

E2E Transport vs. HBH Error Control

• Alternatives
  N none
  O open loop (FEC)
  C closed loop (ARQ)
    • S&W, GB-N, SelRep

• Location
  – HBH
  – E2E

• App requirements
  – unreliable
  – quasi-reliable
  – reliable
Resilient Transport: ResTP

Transmission Control

- Transmission control modes [future work]
  - subflow congestion control
  - subflows should generally not share nodes nor links
ResiliNets Protocols
Cross-Layer Model: ResTP/GeoDivRP

- **Application**
  - \( K_{7 	o 4} = \{ss, tm\} \)
    service spec and threat model

- **E2E Transport: ResTP**
  - erasure spreading vs. hot standby
  - FEC vs. HARQ vs. ARQ
  - \( K_{4 	o 3} = \{k, d, [h, t]\} \)
    \( k \)-path diversity over distance \( d \)
    opt. stretch \( h \) and skew \( t \) bounds

- **Routing: GeoDivRP**
  - construct \( k, d \)-diverse paths
Geodiverse Routing Protocol
GeoDivRP using iWPSP and MLW

• Two heuristics: iWPSP and MLW
• iWPSP (iterative waypoint shortest path)
  – choose neighbours and waypoints to meet diversity spec
  – splice Dijkstra shortest paths
  – complexity: $2c^2n^2 \log n$ (for average of $c$ neighbours)
  – [Cheng and Sterbenz @ KU: DRCN 2014]
• MLW (modified link weights)
  – modify link weights higher close to primary path
  – forces (weighted) shortest path alternates to be diverse
  – complexity: $2n \log n$
  – [Gardner, May, and Medhi @ UMKC: DRCN 2014]
Geodiverse Multipath Routing

GeoDivRP: iWPSP

- GeoDivRP: intermediate waypoint algorithm
  - LSAs contain geolocation of routers
GeoDivRP: intermediate waypoint algorithm
- LSAs contain geolocation of routers
- choose $k$ next hop routers at least $d$ apart if possible
GeoDiverse Multipath Routing

**GeoDivRP: iWPSP**

- **GeoDivRP**: intermediate waypoint algorithm
  - LSAs contain geolocation of routers
  - choose $k$ next hop routers at least $d$ apart if possible
  - choose mid-point waypoints $d + \delta$ wrt to shortest path
    - limit stretch to $h$ and skew to $t$ if specified and possible
• GeoDivRP: intermediate waypoint algorithm
  – LSAs contain geolocation of routers
  – choose $k$ next hop routers at least $d$ apart if possible
  – choose mid-point waypoints $d + \delta$ wrt to shortest path
    • limit stretch to $h$ and skew to $t$ if specified and possible
  – use conventional SPF (Dijkstra) for paths to waypoints
• GeoDivRP: intermediate waypoint algorithm
  – LSAs contain geolocation of routers
  – choose $k$ next hop routers at least $d$ apart if possible
  – choose mid-point waypoints $d + \delta$ wrt to shortest path
    • limit stretch to $h$ and skew to $t$ if specified and possible
  – use conventional SPF (Dijkstra) for paths to waypoints
• GeoDivRP: intermediate waypoint algorithm
  – LSAs contain geolocation of routers
GeoDiverse Multipath Routing

GeoDivRP: MLW

- GeoDivRP: intermediate waypoint algorithm
  - LSAs contain geolocation of routers
  - choose primary shortest path
• GeoDivRP: intermediate waypoint algorithm
  – LSAs contain geolocation of routers
  – choose primary shortest path
  – modify link weights higher close to primary path
  – forces (weighted) shortest path alternates to be diverse
End