Ontology Based Information Centric Tactical Edge Networking

Joseph B. Kopena[†]
Thomas A. Wambold
Christopher T. Cannon
Duc Nguyen
Marcello Balduccini
Emily C. LeBlanc
William C. Regli

[†]tjkopena@bellerophonmobile.com

February 12, 2015

Drexel UNIVERSITY

BELLEROPHON MOBILE

Bellerophon Mobile:

Research and development in computer networking and information management for mobile & extreme environments

Startup out of the Applied Informatics Group at Drexel U

- Extensive history with NSF and DoD applied research
  - Over 10 years' experience in networking and AI projects for DARPA, CERDEC, ONR, NRL, DISA

Four core competency areas

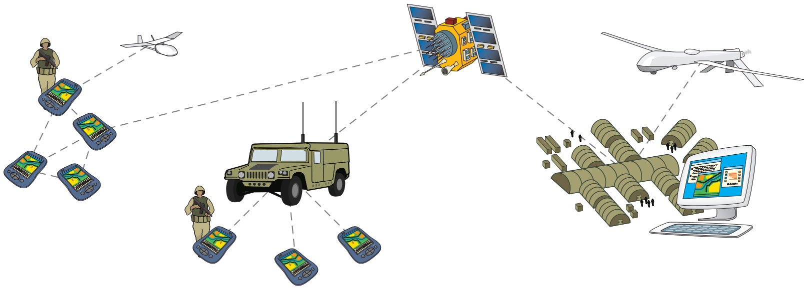
- **Information Management**: Knowledge representation, automated reasoning, and the Semantic Web for content curation & discovery
- **Mobile Networking**: Ad hoc, content-based, and delay tolerant protocols at network, transport, and application layers
- **Network Testing and Evalution**: Experimentation on mobile networks in emulation, simulation, and live testing
- **Mobile App Development**: Android and HTML5 apps, including consumer facing products for Fortune 500 companies

CBMEN THOR

Reporting general experiences from just-concluded DARPA Content-Based Mobile Edge Networking (CBMEN) project

Drexel & Bellerophon responsible for content naming, search, discovery:

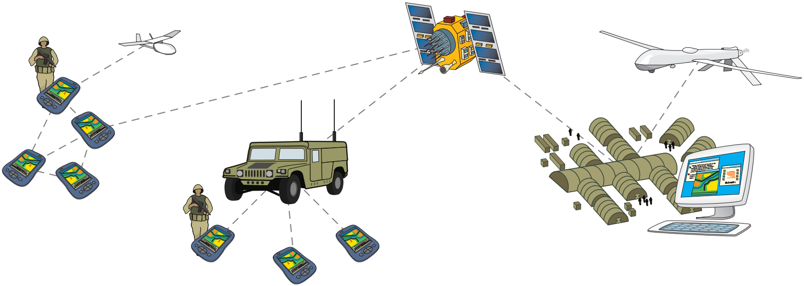
THOR: Tactical Heterogeneous Ontology Representation



Battlefield networks are already prevalent and yet severely limited

- Semi-autonomous groups of mobile users, constrained backbone links
- Severe latency, disconnection, disruption are intrinsic domain features
- Wide variety of devices, networks, apps, data, policies
- Lots of digital data already being generated
- Mostly exchanged only between missions and to/from command

Changing Challenges

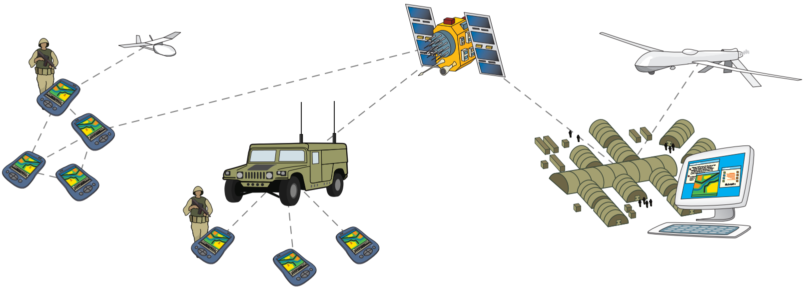


Historically the dominant issues were impoverished nodes and links

Real issues right now are arguably disconnectivity and network scale

- Nodes have significant memory and processing resources
- Network shape is complex, can have substantial link bandwidth
  - High capacity at command & in edge groups, limited capacity between
- Link disruption and network partitioning are pervasive though
- Scaling from handfuls of nodes to units to battlefield is a challenge

Upcoming Problems



Maturing and improving networks will raise challenges new to this domain

Near-future critical issue is information management

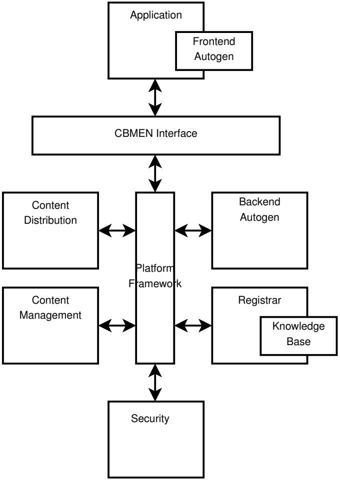
- Better systems will result in network and users deluged with data
  - Have to make good decisions about routing and presenting content
  - Have to discover relevant content, not just return specific search results

CBMEN addresses disconnects and limited bandwidth via information centric networking

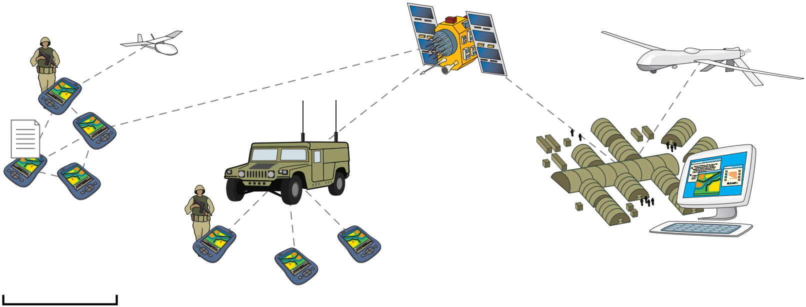
- Content itself is the primary network addressable element
  - Not node endpoints, e.g., servers/clients
- Essentially every node acts as a cache and may provide requested content

THOR enhances this and prepares for information management challenge by addressing and finding content via expresive, ontology-based metadata

- Applies Semantic Web technologies to tactical edge networking



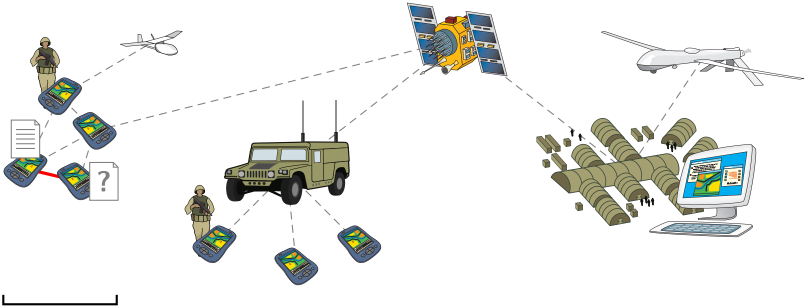
Expressive pub/sub & universal caching enable efficiency, robustness



An application produces a message reporting observed activity

- Local peer node has a matching, long-lived interest and receives report
- Remote node also has interest, but reachback links constrained
  - Detailed matching of content eliminates unnecessary transfers
  - *"All SpotReport messages of activity A in region (L,T)–(R,B)."*
- Later ad hoc query resolved with cached content, obviating transfers

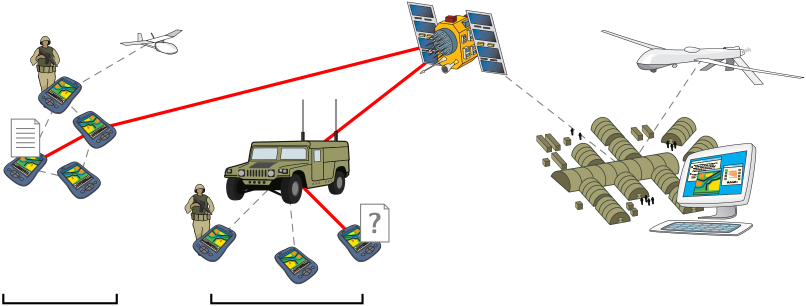
Expressive pub/sub & universal caching enable efficiency, robustness



An application produces a message reporting observed activity

- Local peer node has a matching, long-lived interest and receives report
- Remote node also has interest, but reachback links constrained
  - Detailed matching of content eliminates unnecessary transfers
  - *"All SpotReport messages of activity A in region (L,T)–(R,B)."*
- Later ad hoc query resolved with cached content, obviating transfers

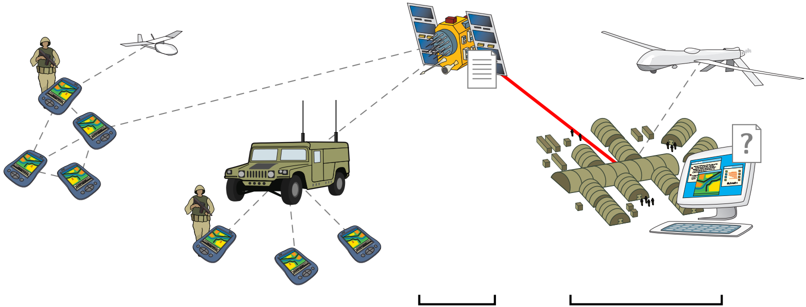
Expressive pub/sub & universal caching enable efficiency, robustness



An application produces a message reporting observed activity

- Local peer node has a matching, long-lived interest and receives report
- Remote node also has interest, but reachback links constrained
  - Detailed matching of content eliminates unnecessary transfers
  - *"All SpotReport messages of activity A in region (L,T)–(R,B)."*
- Later ad hoc query resolved with cached content, obviating transfers

Expressive pub/sub & universal caching enable efficiency, robustness



An application produces a message reporting observed activity

- Local peer node has a matching, long-lived interest and receives report
- Remote node also has interest, but reachback links constrained
  - Detailed matching of content eliminates unnecessary transfers
  - *"All SpotReport messages of activity A in region (L,T)–(R,B)."*
- Later ad hoc query resolved with cached content, obviating transfers

Bellerophon Mobile

CBMEN content is described in RDF

- Generated by applications directly, or through shared Autogen user interface
- Generally one piece of metadata per piece of content, but it's a true generic knowledge base and structures can be built connecting content or completely independent data

```
<urn:registrar:mc#c5cb82...a2aa>
  a messages:SpotReport;
  messages:contentType [ a provenance:ImageEntity ];
  messages:contentFormat [ a messages:JPG ];

  c2:latitude "38.165048" ;
  c2:longitude "-77.284355" ;

  rdfs:comment "From Olivia" .
```

Searches and proactive subscriptions are specified as SPARQL queries

OWL-Lite ontologies define a few standard structures and substantial taxonomies for the domain that are applied in matching

```
SELECT ?id ?format ?comment
WHERE {
  ?id a messages:SpotReport;
      messages:contentType ?ct .
  ?ct a provenance:ImageEntity .
  OPTIONAL { ?id messages:contentFormat ?cf .
             ?cf a ?format .
             MINUS { ?cf a ?cfsub .
                     ?cfsub rdfs:subClassOf ?format .
             }
  }
  OPTIONAL { ?id rdfs:comment ?comment . }
}
```

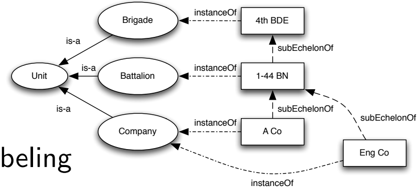
Metadata Inference

Bellerophon Mobile

Semantic Web metadata supports both specificity and generalization

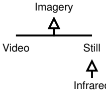
RDF metadata enables complex modeling of content descriptions and queries

- Georeferencing, units, roles, content summarization tags, etc.
- Impractical with flat or hierarchical labeling

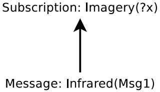


OWL-Lite enables just enough inference

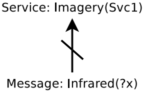
- Apply implicit knowledge from background ontologies (echelons, etc.)
- Derive implications and connections from context (task org, roles, etc.)



Subclass relationships in notional hierarchy of imagery producer or consumer types



A specific publication will be delivered to a more general subscriber



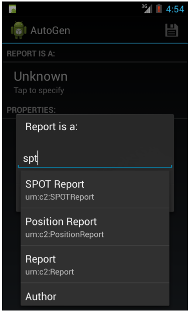
A specific request will not be matched to a more general search

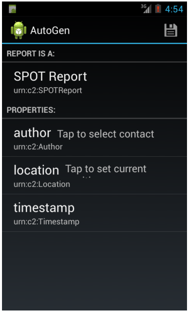
Markup

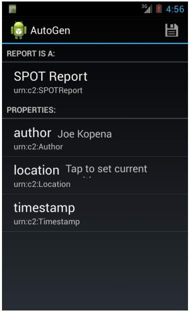
Current demo apps generate metadata using THOR's Autogen component

Autogen: An Android activity any app can invoke to construct RDF

- Caller specifies a base class and/or partially populated structure
- Interface queries ontologies for subclasses and fields, populating the interface and presenting autocomplete options as user types
- Literal data fields may invoke other activities for input
  - Contacts for person, GPS or map for coordinate, calendar for date, etc.

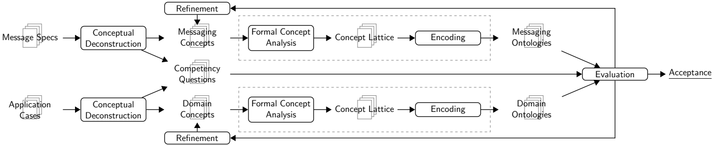






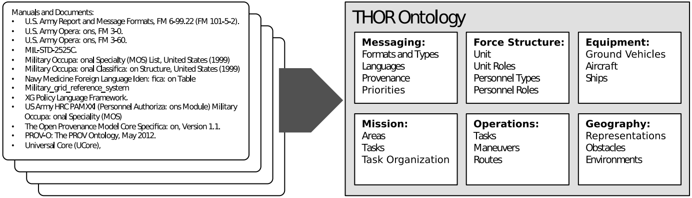
Ontology

The THOR ontology constructed by distilling doctrine and canonical apps



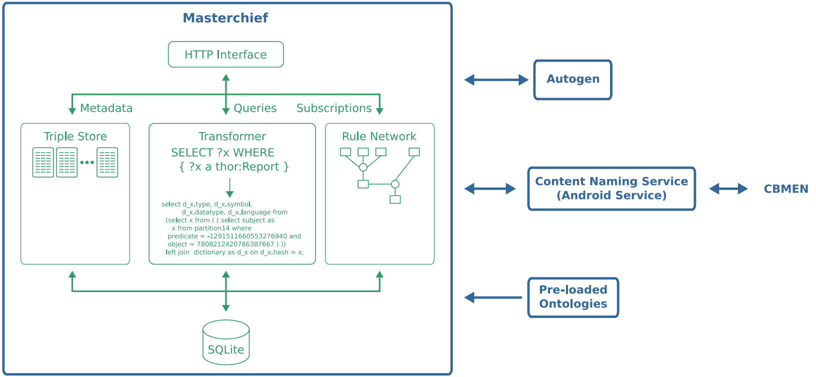
Several core structures and extensive taxonomies, organized into modules

- Structures: Messages, force organization, missions, provenance, etc.
- Taxonomies: Nations, unit types, equipment, observations, tasks, etc.
- End product is ~3700 classes, ~100 properties



Search and subscriptions are executed on all nodes by Masterchief

- Mobile-ready knowledge base built for CBMEN
- Semantic Web logic implemented over SQLite for robustness; portability; and limited, constant-size primary memory utilization
- Transformations, management, and interface in C & Go for portability



Lessons

A little semantics goes a long way

- Potential stakeholders primarily interested in basic taxonomies

Fairly difficult to get developers without KR experience up to speed

- Project apps didn't get to point of utilizing capabilities for collaboration, versioning, etc., offered by the underlying model

Evaluation of KR systems is extremely difficult

- Performance is non-trivial but fairly straightforward
  - Sidenote: What's hard for network may not be hard for KR, & vice versa
- But testing actual effectiveness and value requires complex yet realistic scenarios, revolves around metrics that are difficult to quantify

SPARQL and RDF model aren't quite the right tools for this task

- SPARQL great for querying the KB, less ideal for fetching objects
  - Apps want all the metadata about content, resulting in massive queries
- RDF+SPARQL cumbersome when working with dynamic data
  - E.g., "All reports within 3 miles of *my current position*."