Supporting Story Synthesis: Bridging the Gap between Visual Analytics and Storytelling

Siming Chen, Jie Li, Gennady Andrienko, Natalia Andrienko, Yun Wang, Phong H. Nguyen, and Cagatay Turkay

(2018)

Background **Research Problem Related Works Bridging the Gap Proposed Solution** Thoughts

Background **Research Problem Related Works Bridging the Gap Proposed Solution** Thoughts

Visual Analytics



Figure 6. Visual analytics interface of social media, including (a) topic projection view, (b) topic comparison view, (c) user/ location distribution view, (d) temporal view, (e) story slice generation parameters, and (f) raw data table.

Visual Analytics



Figure 6. Visual analytics interface of social media, including (a) topic projection view, (b) topic comparison view, (c) user/ location distribution view, (d) temporal view, (e) story slice generation parameters, and (f) raw data table.

Storytelling



Figure 8. Layout methods to organize story slices according to different perspectives: (a) time cycles, (b) locations, (c) users, and (d) keywords.

Storytelling



Figure 8. Layout methods to organize story slices according to different perspectives: (a) time cycles, (b) locations, (c) users, and (d) keywords.

Storytelling

- Helps non-experts connect the dots
- Effectively communicates insights
- Provides meaning and context to data

Background **Research Problem Related Works Bridging the Gap Proposed Solution** Thoughts













Background **Research Problem Related Works Bridging the Gap Proposed Solution** Thoughts

Past Work

- Need for organizing story contents
- Existing works compose stories with annotated screen shots of analytic displays
- Provenance systems rely on completed analysis bookmarks
- Other systems rely on completed visualizations

Past Work

Need for organizing story contents

Inherits visual complexity of analytic displays

- Existing works compose stories with annotated screen shots of analytic displays
- Provenance systems rely on completed analysis bookmarks
- Other systems rely on completed visualizations

Past Work

Need for organizing story contents

Inherits visual complexity of analytic displays

- Existing works compose stories with annotated screen shots of analytic displays
 No current systems
- Provenance systems rely on comple bookmarks
 are integrated into the analysis workflow!
- Other systems rely on completed visualizations

Background **Research Problem Related Works** Bridging the Gap **Proposed Solution** Thoughts



Figure 6. Proposed framework for bridging the gap.



Figure 6. Proposed framework for bridging the gap.



Figure 6. Proposed framework for bridging the gap.



Figure 6. Proposed framework for bridging the gap.

- Structured representations of findings during the course of analysis
- Define what facets (information) to capture

VAST Challenge 2011 finding structure ::=

<label, time, location, N people, N messages, {(keyword, frequency)}, context>

- Structured representations of findings during the course of analysis
- Define what facets (information) to capture

VAST Challenge 2011 finding structure ::=

<label, time, location, N people, N messages, {(keyword, frequency)}, context>

VAST Challenge, analyze the circumstances of an epidemic outbreak via microblogs

- Structured representations of findings during the course of analysis
- Define what facets (information) to capture

VAST Challenge 2011 finding structure ::=



Analytic Phase

Produce story slices during Analyze tasks



Figure 5. Visual analytics phase. After topic modelling, analysts explore the collection of messages from four perspectives and extract story slices.

<F1 (outbreak), time = May 18-20, location = Vastopolis, N people = 27 446, N messages = 59 755
<F2 (cluster center-east), time = May 18-20, location = polygon1, N people = 16 479, N messages
<F3 (cluster southwest), time = May 19-20, location = polygon2, N people = 6 752, N messages =
<F4 (hospitals), time = May 20, location = {(-93.33, 42.24), (-93.42, 42.25), (-93.44, 42.20), ...}, N p
<f5 (truck accident), time=May 17, location = (-93.427,42.226), N people = 149, N messages=149,</pre>

59 755, {(chills, 10 436), (fever, 7 585), (sick, 6 543), ...}, null>

ages = 32 445, {(chills, 6 511), (fever, 4 905), ..., (flu, 3 466), ...}, context = {wind = west-to-east}>

ges = 9 719, {(diarrhea, 2 785), (stomach, 2 682), ..., (nausea, 766), ...}, context = {river flow = north-to-southwest}>

.}, N people = 3 265, N messages = 3 276, {{(chills, 1 419), (fever, 1 171), ..., (flu, 886), ...}, null>

=149, {(truck, 127), ..., (accident, 37), ..., (burning, 14), ..., (spilling cargo, 9), ...}, context = {motorway, river bridge}>



<F1 (outbreak), time = May 18-20, location = Vastopolis, N people = 27 446, N messages = 59 755
<F2 (cluster center-east), time = May 18-20, location = polygon1, N people = 16 479, N messages
<F3 (cluster southwest), time = May 19-20, location = polygon2, N people = 6 752, N messages =
<F4 (hospitals), time = May 20, location = {(-93.33, 42.24), (-93.42, 42.25), (-93.44, 42.20), ...}, N p
<f5 (truck accident), time=May 17, location = (-93.427,42.226), N people = 149, N messages=149, N</pre>

59 755, {(chills, 10 436), (fever, 7 585), (sick, 6 543), ...}, null>

sages = 32 445, {(chills, 6 511), (fever, 4 905), ..., (flu, 3 466), ...}, context = {wind = west-to-east}> ges = 9 719, {(diarrhea, 2 785), (stomach, 2 682), ..., (nausea, 766), ...}, context = {river flow = north-to-southwest}>

- In the second second
- =149, {(truck, 127), ..., (accident, 37), ..., (burning, 14), ..., (spilling cargo, 9), ...}, context = {motorway, river bridge}>



<F1 (outbreak), time = May 18-20, location = Vastopolis, N people = 27 446, N messages = 59 755 <F2 (cluster center-east), time = May 18-20, location = *polygon1*, N people = 16 479, N messages <F3 (cluster southwest), time = May 19-20, location = *polygon2*, N people = 6 752, N messages = <F4 (hospitals), time = May 20, location = {(-93.33, 42.24), (-93.42, 42.25), (-93.44, 42.20), ...}, N p <f5 (truck accident), time=May 17, location = (-93.427,42.226), N people = 149, N messages=149,

59 755, {(chills, 10 436), (fever, 7 585), (sick, 6 543), ...}, null>

sages = 32 445, {(chills, 6 511), (fever, 4 905), ..., (flu, 3 466), ...}, context = {wind = west-to-east}> ges = 9 719, {(diarrhea, 2 785), (stomach, 2 682), ..., (nausea, 766), ...}, context = {river flow = north-to-southwest}> .}, N people = 3 265, N messages = 3 276, {{(chills, 1 419), (fever, 1 171), ..., (flu, 886), ...}, null>

=149, {(truck, 127), ..., (accident, 37), ..., (burning, 14), ..., (spilling cargo, 9), ...}, context = {motorway, river bridge}>



Facet Types

- Discrete categories with no relationships
- Discrete entities with domain-specific relationships
- Linearly ordered elements
- Linearly ordered elements with distances
- ► Time
- ID/2D/3D space

Facet Types

- Discrete categories with no relationships
- Discrete entities with domain-specific relationships
- Linearly ordered elements
- Linearly ordered elements with distances
- ► Time
- ID/2D/3D space

Facets specify possible arrangements and aggregations

Facet Types

- Discrete categories with no relationships
- Discrete entities with domain-specific relationships
- Linearly ordered elements
- Linearly ordered elements with distances
- Time
 1D/2D/3D space
 These are very similar to the Abstract Data Types in VAD...
 s specify possible gements and aggregations

Background **Research Problem Related Works Bridging the Gap Proposed Solution** Thoughts

Story synthesis creates story contents from story slices by...

- Aggregating and summarizing story slices according to facets/measurements
- Embedding details into views/layouts
- Arranging story slices onto layouts according to facets/ relationships
- Showing facets by priority and from multiple perspectives
- Annotating

Story synthesis creates story contents from story slices by...

- Aggregating and summarizing story slices according to facets/measurements
- Embedding details into views/layouts
- Arranging story slices onto layouts according to facets/ relationships
- Showing facets by priority and from multiple perspectives
- Annotating

Story synthesis does not include the design of the final story!









New Framework

- 1. Define **types** and **structures** of story slices What facets or patterns may be used from the data analysis phase?
- 2. Design **representations** of story slices What data structure will be used to contain slice information? How will slices be visualized in the synthesis workspace?
- 3. Define story synthesis support functions.
- Design the visual analytics system, including support for discovery of potential story slices.

New Framework

- 1. Define **types** and **structures** of story slices What facets or patterns may be used from the data analysis phase?
- 2. Design **representations** of story slices What data structure will be used to contain slice information? How will slices be visualized in the synthesis workspace?
- 3. Define story synthesis support fun
- Design the visual analytics system, used in designing visual discovery of potential story slices. analytics systems that

This framework can be used in designing visual analytics systems that provide support for story synthesis.

Background **Research Problem Related Works Bridging the Gap Proposed Solution** Thoughts

Good

• Need for framework is well backed by current research

Story slices help reason about abstract data more

Story slices capture model of analysis concept generation

Evaluated by two sociology experts (positive feedback)

Bad

O Paper didn't do a great job proving generality of framework

ONAJORITY OF EXAMPLES USED ONLINE TEXT MESSAGE DATA

O Paper structure could be improved to help understanding

Solution Is this framework better suited for messaging data?