

Going Beyond Corr-LDA for Detecting Specific Comments on News & Blogs

Trapit Bansal

Joint work with Mrinal Das & Prof. Chiranjib Bhattacharyya

Machine Learning Lab
Department of CSA, IISc

16th January, 2014

Understanding Comments

YAHOO! NEWS

News Home
U.S.
World
Politics
Tech
Science
Health
Cold News
Opinion
Local
Dear Abby
Comics
ABC News
Yahoo Originals
Photos

Related
Robert Gates history...

Search News [Advanced Search](#)

Obama: 2014 can be 'breakthrough year for America'

AP By JARIE RACE November 20, 2013 4:08 PM



WASHINGTON (AP) — Citing strong progress on the economy, President Barack Obama said at his annual year-end news conference Friday that 2014 "can be a breakthrough year for America" after a long season of recession and slow recovery.

Yet he suggested that given widespread criticism, he may alter the power of the National Security Agency to collect information on Americans.

14,972 Comments

Sign in to post a comment

Popular New Oldest Most Replied

Jim · 1 hour ago
Yahoo!ARConey insulting our intelligence by keeping a bad story up for nearly a month. There is so much going on, including severe repercussions by Fox that RFD was ordered by Twitter and 200 Best Blogger news a month ago while it was underway. It is an abuse of the 1st Amendment to keep this page propagandize up. RFD had for several weeks, to ensure he could be re-elected. Had that been Bush, Yahoo would be calling for his impeachment instead.
More
Report Abuse (0) Reply

Debra · 1 hour ago
Obama lies to you daily.
In October of last year, obama's health care agency, gave a half-million-dollar grant to a registered lobbying firm to help enroll people for Obamacare as Affordable Care Act navigators. The Centers for Medicare and Medicaid Services, gave the \$510,577 grant to Chatham LLC, which has tied disclosure.
More
Reply

Most Interesting Troll In The World · 1 hour ago
Obama's vision for 2014 is Obamacare. It's best explained by example:
Socialism: You have two cows. The government takes one and gives it to your neighbor.
Communism: You have two cows. You give them to the Government, and the Government then gives you some milk.
More
Reply

Abortion Is Murder! · 1 hour ago
There are 18 Democratic Senators running for election in 2014 who know you'll lose your health insurance because of Obama-care, but promised you could keep it anyway.
These are the LINGS who STOLE your healthcare!
1) Mark Begich (Alaska)
More
Reply

[View Thread Statistics](#)

- 15,000 comments on a single article
- Most research mostly focused on detecting spam
- There are no models which could understand comments and relate them to the source, such as the article.

Types of Comments

Cheaper than most, better than all: the 2013 Nexus 7 reviewed

Just over a year ago, Google released its first Nexus tablet . The 2012 Nexus 7 wasn't perfect by a long shot, but it was the kick in the pants that the Android tablet ecosystem needed at the time.

.....

The 1920x1200 display has much brighter colors than last year's 1280x800 panel, and at 323 PPI, it outdoes both the 300 PPI Nexus 10 and 264 PPI Retina iPads. Until there's a Retina iPad mini, no other small-tablet screen comes close.

.....

In our previous article, we noted that the 2013 Nexus 7 was consistently faster than the Nexus 4 despite the fact that the two ostensibly share the same system-on-a-chip (SoC)

.....

If Apple responds with a Retina-equipped iPad mini in the fall, the balance of power may shift back in the other direction. But if it sticks with its current display, it will become more difficult to recommend.

Comments

As always, an authoritative review. I picked one up after getting tired of Apple's iPad Mini Retina Display nonsense -they want to prolong the agony before releasing it and eating away their iPad revenue. I need the high resolution screen because....

It's good to see mainstream tablets are getting cheaper with better technology.I just got a 10 inch Chinese tablet for about the same as this 7 in mainstream one.

*I agree, but your very example just proved that apple *didn't* care about the user experience and just wanted to release an ipad with a retina screen in the early part of 2012 without regard for the "experience".*

Yup.I read, play games, do email, check quick facts and goof off on my tablet. I surf on a laptop. I work on a laptop.

Types of Comments

Cheaper than most, better than all: the 2013 Nexus 7 reviewed

Just over a year ago, Google released its first Nexus tablet . The 2012 Nexus 7 wasn't perfect by a long shot, but it was the kick in the pants that the Android tablet ecosystem needed at the time.

The 1920x1200 display has much brighter colors than last year's 1280x800 panel, and at 323 PPI, it outdoes both the 300 PPI Nexus 10 and 264 PPI Retina iPads. Until there's a Retina iPad mini, no other small-tablet screen comes close.

In our previous article, we noted that the 2013 Nexus 7 was consistently faster than the Nexus 4 despite the fact that the two ostensibly share the same system-on-a-chip (SoC)

If Apple responds with a Retina-equipped iPad mini in the fall, the balance of power may shift back in the other direction. But if it sticks with its current display, it will become more difficult to recommend.

Comments

As always, an authoritative review. I picked one up after getting tired of Apple's iPad Mini Retina Display nonsense -they want to prolong the agony before releasing it and eating away their iPad revenue. I need the high resolution screen because....

It's good to see mainstream tablets are getting cheaper with better technology.I just got a 10 inch Chinese tablet for about the same as this 7 in mainstream one.

*I agree, but your very example just proved that apple *didn't* care about the user experience and just wanted to release an ipad with a retina screen in the early part of 2012 without regard for the "experience".*

Yup.I read, play games, do email, check quick facts and goof off on my tablet. I surf on a laptop. I work on a laptop.

 **Specific**

Types of Comments

Cheaper than most, better than all: the 2013 Nexus 7 reviewed

Just over a year ago, Google released its first Nexus tablet . The 2012 Nexus 7 wasn't perfect by a long shot, but it was the kick in the pants that the Android tablet ecosystem needed at the time.

The 1920x1200 display has much brighter colors than last year's 1280x800 panel, and at 323 PPI, it outdoes both the 300 PPI Nexus 10 and 264 PPI Retina iPads. Until there's a Retina iPad mini, no other small-tablet screen comes close.

In our previous article, we noted that the 2013 Nexus 7 was consistently faster than the Nexus 4 despite the fact that the two ostensibly share the same system-on-a-chip (SoC)

If Apple responds with a Retina-equipped iPad mini in the fall, the balance of power may shift back in the other direction. But if it sticks with its current display, it will become more difficult to recommend.

Comments

As always, an authoritative review. I picked one up after getting tired of Apple's iPad Mini Retina Display nonsense -they want to prolong the agony before releasing it and eating away their iPad revenue. I need the high resolution screen because....

It's good to see mainstream tablets are getting cheaper with better technology.I just got a 10 inch Chinese tablet for about the same as this 7 in mainstream one.

*I agree, but your very example just proved that apple *didn't* care about the user experience and just wanted to release an ipad with a retina screen in the early part of 2012 without regard for the "experience".*

Yup.I read, play games, do email, check quick facts and goof off on my tablet. I surf on a laptop. I work on a laptop.

— Specific

Specific comments

Comments are related to **specific portion of the article**

Types of Comments

Cheaper than most, better than all: the 2013 Nexus 7 reviewed

Just over a year ago, Google released its first Nexus tablet . The 2012 Nexus 7 wasn't perfect by a long shot, but it was the kick in the pants that the Android tablet ecosystem needed at the time.

.....

The 1920x1200 display has much brighter colors than last year's 1280x800 panel, and at 323 PPI, it outdoes both the 300 PPI Nexus 10 and 264 PPI Retina iPads. Until there's a Retina iPad mini, no other small-tablet screen comes close.

.....

In our previous article, we noted that the 2013 Nexus 7 was consistently faster than the Nexus 4 despite the fact that the two ostensibly share the same system-on-a-chip (SoC)

.....

If Apple responds with a Retina-equipped iPad mini in the fall, the balance of power may shift back in the other direction. But if it sticks with its current display, it will become more difficult to recommend.

Comments

As always, an authoritative review. I picked one up after getting tired of Apple's iPad Mini Retina Display nonsense -they want to prolong the agony before releasing it and eating away their iPad revenue. I need the high resolution screen because....

It's good to see mainstream tablets are getting cheaper with better technology.I just got a 10 inch Chinese tablet for about the same as this 7 in mainstream one.

I agree, but your very example just proved that apple *didn't* care about the user experience and just wanted to release an ipad with a retina screen in the early part of 2012 without regard for the "experience".

Yup.I read, play games, do email, check quick facts and goof off on my tablet. I surf on a laptop. I work on a laptop.

—— Specific

—— General

Types of Comments

Cheaper than most, better than all: the 2013 Nexus 7 reviewed

Just over a year ago, Google released its first Nexus tablet . The 2012 Nexus 7 wasn't perfect by a long shot, but it was the kick in the pants that the Android tablet ecosystem needed at the time.

.....

The 1920x1200 display has much brighter colors than last year's 1280x800 panel, and at 323 PPI, it outdoes both the 300 PPI Nexus 10 and 264 PPI Retina iPads. Until there's a Retina iPad mini, no other small-tablet screen comes close.

.....

In our previous article, we noted that the 2013 Nexus 7 was consistently faster than the Nexus 4 despite the fact that the two ostensibly share the same system-on-a-chip (SoC)

.....

If Apple responds with a Retina-equipped iPad mini in the fall, the balance of power may shift back in the other direction. But if it sticks with its current display, it will become more difficult to recommend.

Comments

As always, an authoritative review. I picked one up after getting tired of Apple's iPad Mini Retina Display nonsense -they want to prolong the agony before releasing it and eating away their iPad revenue. I need the high resolution screen because....

It's good to see mainstream tablets are getting cheaper with better technology.I just got a 10 inch Chinese tablet for about the same as this 7 in mainstream one.

*I agree, but your very example just proved that apple *didn't* care about the user experience and just wanted to release an ipad with a retina screen in the early part of 2012 without regard for the "experience".*

Yup.I read, play games, do email, check quick facts and goof off on my tablet. I surf on a laptop. I work on a laptop.

—— Specific

—— General

General comments

Comments related to the entire article

Types of Comments

Cheaper than most, better than all: the 2013 Nexus 7 reviewed

Just over a year ago, Google released its first Nexus tablet . The 2012 Nexus 7 wasn't perfect by a long shot, but it was the kick in the pants that the Android tablet ecosystem needed at the time.

.....

The 1920x1200 display has much brighter colors than last year's 1280x800 panel, and at 323 PPI, it outdoes both the 300 PPI Nexus 10 and 264 PPI Retina iPads. Until there's a Retina iPad mini, no other small-tablet screen comes close.

.....

In our previous article, we noted that the 2013 Nexus 7 was consistently faster than the Nexus 4 despite the fact that the two ostensibly share the same system-on-a-chip (SoC)

.....

If Apple responds with a Retina-equipped iPad mini in the fall, the balance of power may shift back in the other direction. But if it sticks with its current display, it will become more difficult to recommend.

Comments

As always, an authoritative review. I picked one up after getting tired of Apple's iPad Mini Retina Display nonsense -they want to prolong the agony before releasing it and eating away their iPad revenue. I need the high resolution screen because....

It's good to see mainstream tablets are getting cheaper with better technology.I just got a 10 inch Chinese tablet for about the same as this 7 in mainstream one.

*I agree, but your very example just proved that apple *didn't* care about the user experience and just wanted to release an ipad with a retina screen in the early part of 2012 without regard for the "experience".*

Yup.I read, play games, do email, check quick facts and goof off on my tablet. I surf on a laptop. I work on a laptop.

Specific

General

Irrelevant

Types of Comments

Cheaper than most, better than all: the 2013 Nexus 7 reviewed

Just over a year ago, Google released its first Nexus tablet . The 2012 Nexus 7 wasn't perfect by a long shot, but it was the kick in the pants that the Android tablet ecosystem needed at the time.

.....

The 1920x1200 display has much brighter colors than last year's 1280x800 panel, and at 323 PPI, it outdoes both the 300 PPI Nexus 10 and 264 PPI Retina iPads. Until there's a Retina iPad mini, no other small-tablet screen comes close.

.....

In our previous article, we noted that the 2013 Nexus 7 was consistently faster than the Nexus 4 despite the fact that the two ostensibly share the same system-on-a-chip (SoC)

.....

If Apple responds with a Retina-equipped iPad mini in the fall, the balance of power may shift back in the other direction. But if it sticks with its current display, it will become more difficult to recommend.

Comments

As always, an authoritative review. I picked one up after getting tired of Apple's iPad Mini Retina Display nonsense -they want to prolong the agony before releasing it and eating away their iPad revenue. I need the high resolution screen because....

It's good to see mainstream tablets are getting cheaper with better technology.I just got a 10 inch Chinese tablet for about the same as this 7 in mainstream one.

*I agree, but your very example just proved that apple *didn't* care about the user experience and just wanted to release an ipad with a retina screen in the early part of 2012 without regard for the "experience".*

Yup.I read, play games, do email, check quick facts and goof off on my tablet. I surf on a laptop. I work on a laptop.

 **Specific**

 **General**

 **Irrelevant**

Irrelevant comments

Comments not related to the article (spam or diversionary)

Different types of Comments

Specific comments

Comments related to **specific portion of the article**

A large fraction, about 50%, of the comments are very specific

*I agree, but your very example just proved that apple *didn't* care about the user experience and just wanted to release an ipad with a retina screen in the early part of 2012 without regard for the "experience".*

General comments: Comments broadly related to the article

It's good to see mainstream tablets are getting cheaper with better technology. I just got a 10 inch Chinese tablet for about the same as this 7 in mainstream one.

Irrelevant comments: Comments not related to the article

Yup. I read, play games, do email, check quick facts and goof off on my tablet. I surf on a laptop. I work on a laptop.

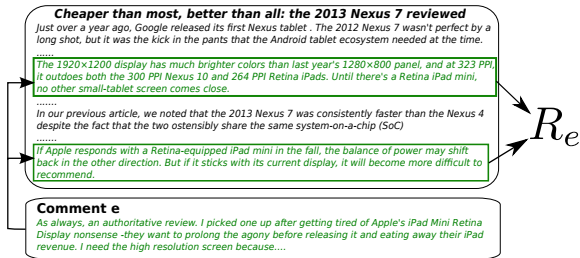
Outline

- 1 Specific Correspondence Location problem
- 2 SCL is hard
- 3 Contributions
- 4 Using SCTM for listening to the user

The SCL problem

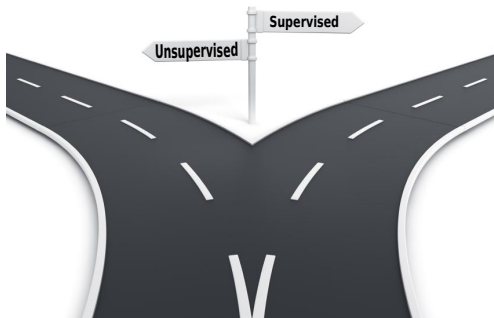
Given a comment locate sentences in the original document which may have motivated the comment. If number of sentences is low, we call it **specific comment**.

Specific Correspondence Location(SCL)

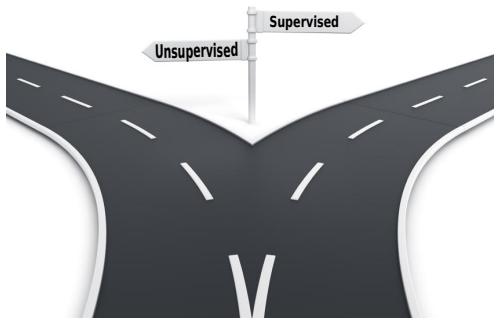


- An article A is a collection of S sentences
- Let $R_e \subseteq A$ denote the set of sentences related to the e -th comment of A
- Comment e is a **specific comment** if $0 < |R_e| \leq N$

Supervised Approaches

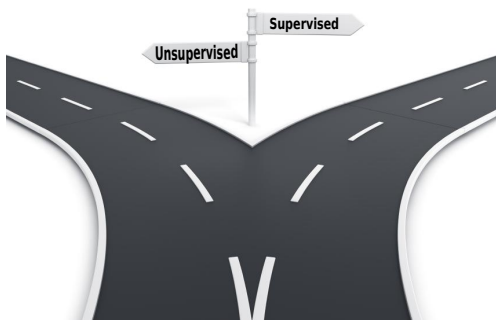


Supervised Approaches



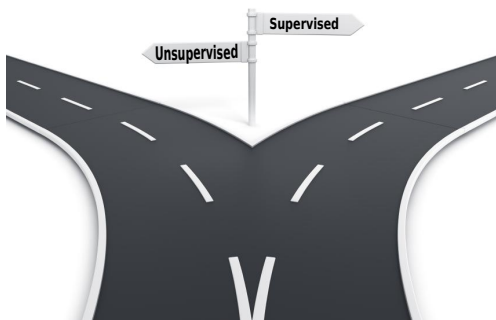
- Lack of labeled data

Supervised Approaches



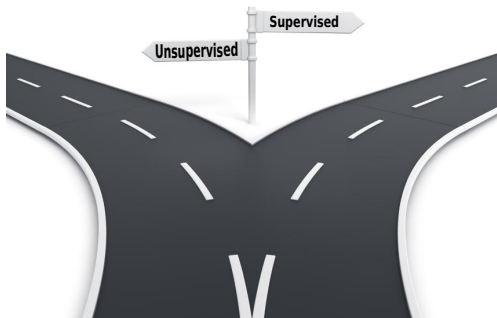
- Lack of labeled data
- Feature engineering remains a challenge - there are no distinguishing features of a **Specific comment**

Supervised Approaches are unsuitable

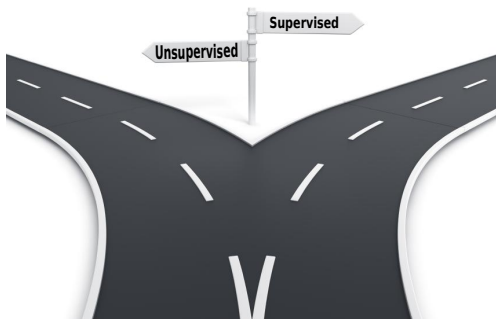


- Lack of labeled data
- Feature engineering remains a challenge - there are no distinguishing features of a **Specific comment**

Unsupervised Approaches:

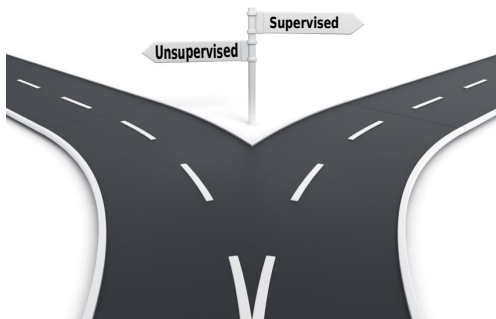


Unsupervised Approaches:



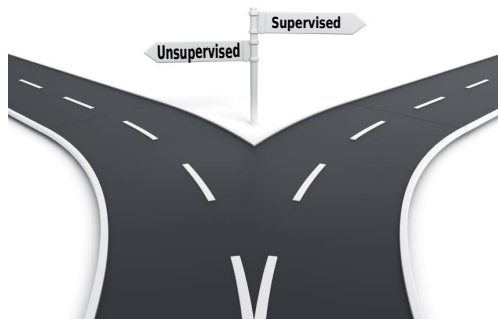
- Topic models [2] have worked well for texts

Unsupervised Approaches:



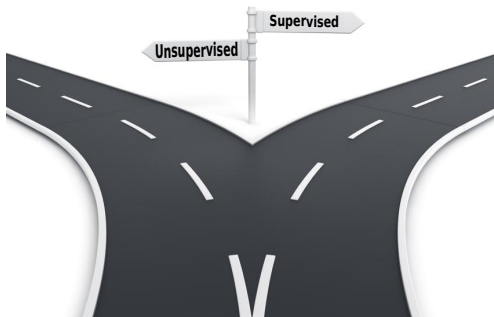
- Topic models [2] have worked well for texts
- Does not require labeled data

Unsupervised Approaches: Suitability of Topic Models



- Topic models [2] have worked well for texts
- Does not require labeled data
- Can use Correspondence-LDA model [1]

Unsupervised Approaches: Suitability of Topic Models

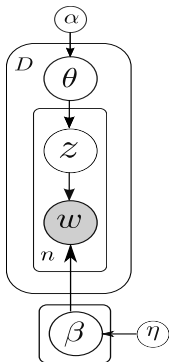


- Topic models [2] have worked well for texts
- Does not require labeled data
- Can use Correspondence-LDA model [1]
It is unsuitable for detecting specific comments

Contributions: Specific Correspondence Topic Model

- Generalize Corr-LDA by introducing Multiple Topic vectors (MTV)
- **Stick breaking process** is a useful prior for MTV
- Develop a **Collapsed Gibbs sampling** approach
- Using SCTM one can detect specific comments. Significant portion of the comments are specific

Latent Dirichlet Allocation – single topic vector



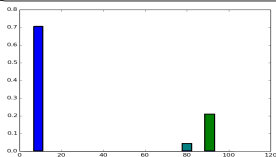
Cheaper than most, better than all: the 2013 Nexus 7 reviewed

Just over a year ago, [Google](#) released its first [Nexus tablet](#). The 2012 [Nexus 7](#) wasn't perfect by a long shot, but it was the kick in the pants that the [Android tablet](#) ecosystem needed at the time.

.....
The short version : The 2013 [Nexus 7](#) is more [compact](#) and, overall, feels a little better put together than last year's [model](#). This is [plastic](#) done right.

....
The 1920x1200 [display](#) has much brighter colors than last year's 1280x800 [panel](#), and at 323 PPI, it outdoes both the 300 PPI [Nexus 10](#) and 264 PPI [Retina iPads](#). Until there's a [Retina iPad mini](#), no other small-tablet [screen](#) comes close.

.....
If [Apple](#) responds with a [Retina-equipped iPad mini](#) in the fall, the balance of power may shift back in the other direction. But if it sticks with its current [display](#), it will become more difficult to recommend.



Topic 10: 0.71
Nexus/Tablets
android:0.027
google:0.018
nexus:0.018
tablet:0.015
phone:0.013
apple:0.010
device:0.008
devices:0.008

Topic 91: 0.2
Phone
phone:0.033
screen:0.013
phones:0.012
android:0.012
samsung:0.010
battery:0.008
galaxy:0.008
iphone:0.007

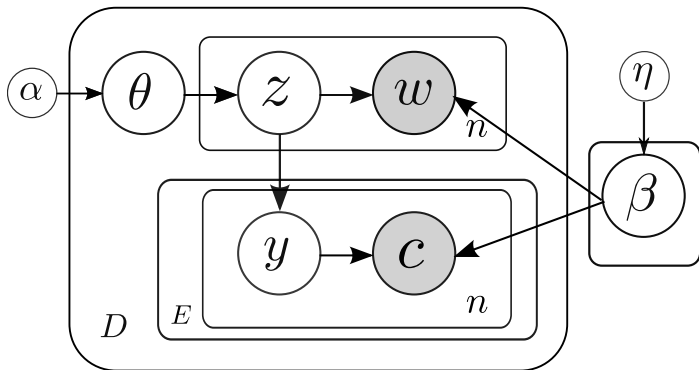
Topics

distribution over the vocabulary

LDA – vanilla topic model

- a topic vector (θ) for a document
- generate the document from the topics, using θ

State of art – Corr-LDA (single topic vector)



- Generative process of article is same as LDA
- Uniformly randomly pick words to get comments

Going beyond single topic vector assumption

- Single topic vector does not allow flexibility in modeling individual sentences
- Need models which can allow variations in topics related to each sentences

Going beyond single topic vector assumption

- Single topic vector does not allow flexibility in modeling individual sentences
- Need models which can allow variations in topics related to each sentences

Key idea: *multiple topic vectors*

use multiple distributions over topics per article. Individual sentences can be drawn from any one of them

Specific Correspondence Topic Model

Cheaper than most, better than all: the 2013 Nexus 7 reviewed

Just over a year ago, Google released its first Nexus tablet. The 2012 Nexus 7 wasn't perfect by a long shot, but it was the kick in the pants that the Android tablet ecosystem needed at the time.

The 1920x1200 display has much brighter colors than last year's 1280x800 panel, and at 323 PPI, it outdoes both the 300 PPI Nexus 10 and 264 PPI Retina iPads. Until there's a Retina iPad mini, no other small-tablet screen comes close.

In our previous article, we noted that the 2013 Nexus 7 was consistently faster than the Nexus 4 despite the fact that the two ostensibly share the same system-on-a-chip (SoC)

If Apple responds with a Retina-equipped iPad mini in the fall, the balance of power may shift back in the other direction. But if it sticks with its current display, it will become more difficult to recommend.

Comments

As always, an authoritative review. I picked one up after getting tired of Apple's iPad Mini Retina Display nonsense -they want to prolong the agony before releasing it and eating away their iPad revenue. I need the high resolution screen because...

It's good to see mainstream tablets are getting cheaper with better technology. I just got a 10 inch Chinese tablet for about the same as this 7 in mainstream one.

I agree, but your very example just proved that apple "didn't" care about the user experience and just wanted to release an ipad with a retina screen in the early part of 2012 without regard for the "experience".

Yup. I read, play games, do email, check quick facts and goof off on my tablet. I surf on a laptop. I work on a laptop.

- Specific comment (green) is related only to the sentences in green box

— Specific

— General

— Irrelevant

Specific Correspondence Topic Model

Cheaper than most, better than all: the 2013 Nexus 7 reviewed

Just over a year ago, Google released its first Nexus tablet. The 2012 Nexus 7 wasn't perfect by a long shot, but it was the kick in the pants that the Android tablet ecosystem needed at the time.

The 1920x1200 display has much brighter colors than last year's 1280x800 panel, and at 323 PPI, it outdoes both the 300 PPI Nexus 10 and 264 PPI Retina iPads. Until there's a Retina iPad mini, no other small-tablet screen comes close.

In our previous article, we noted that the 2013 Nexus 7 was consistently faster than the Nexus 4 despite the fact that the two ostensibly share the same system-on-a-chip (SoC)

If Apple responds with a Retina-equipped iPad mini in the fall, the balance of power may shift back in the other direction. But if it sticks with its current display, it will become more difficult to recommend.

Comments

As always, an authoritative review. I picked one up after getting tired of Apple's iPad Mini Retina Display nonsense -they want to prolong the agony before releasing it and eating away their iPad revenue. I need the high resolution screen because...

It's good to see mainstream tablets are getting cheaper with better technology. I just got a 10 inch Chinese tablet for about the same as this 7 in mainstream one.

I agree, but your very example just proved that apple *didn't* care about the user experience and just wanted to release an ipad with a retina screen in the early part of 2012 without regard for the "experience".

Yup. I read, play games, do email, check quick facts and goof off on my tablet. I surf on a laptop. I work on a laptop.

— Specific — General — Irrelevant

- **Specific comment** (green) is related only to the sentences in green box
- Through multiple topic vectors the green colored sentences in the article can have different distribution over topics than the entire article

Specific Correspondence Topic Model

Cheaper than most, better than all: the 2013 Nexus 7 reviewed

Just over a year ago, Google released its first Nexus tablet. The 2012 Nexus 7 wasn't perfect by a long shot, but it was the kick in the pants that the Android tablet ecosystem needed at the time.

The 1920x1200 display has much brighter colors than last year's 1280x800 panel, and at 323 PPI, it outdoes both the 300 PPI Nexus 10 and 264 PPI Retina iPads. Until there's a Retina iPad mini, no other small-tablet screen comes close.

In our previous article, we noted that the 2013 Nexus 7 was consistently faster than the Nexus 4 despite the fact that the two ostensibly share the same system-on-a-chip (SoC)

If Apple responds with a Retina-equipped iPad mini in the fall, the balance of power may shift back in the other direction. But if it sticks with its current display, it will become more difficult to recommend.

Comments

As always, an authoritative review. I picked one up after getting tired of Apple's iPad Mini Retina Display nonsense -they want to prolong the agony before releasing it and eating away their iPad revenue. I need the high resolution screen because...

It's good to see mainstream tablets are getting cheaper with better technology. I just got a 10 inch Chinese tablet for about the same as this 7 in mainstream one.

I agree, but your very example just proved that apple *didn't* care about the user experience and just wanted to release an ipad with a retina screen in the early part of 2012 without regard for the "experience".

Yup. I read, play games, do email, check quick facts and goof off on my tablet. I surf on a laptop. I work on a laptop.

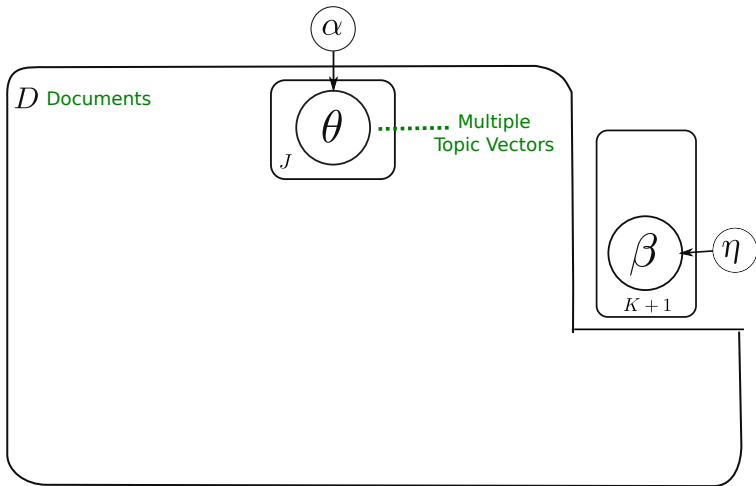
Specific

General

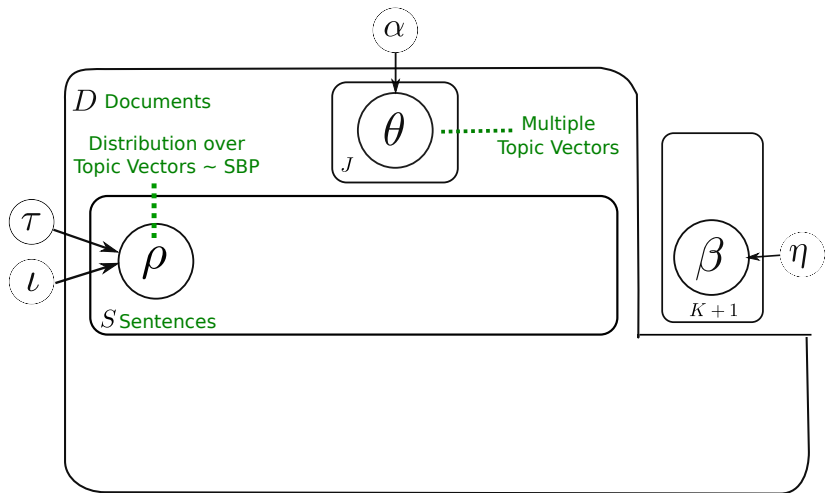
Irrelevant

- Specific comment (green) is related only to the sentences in green box
- Through multiple topic vectors the green colored sentences in the article can have different distribution over topics than the entire article
- Through selection of subset of sentences, the specific comment can pick its topics from these sentences

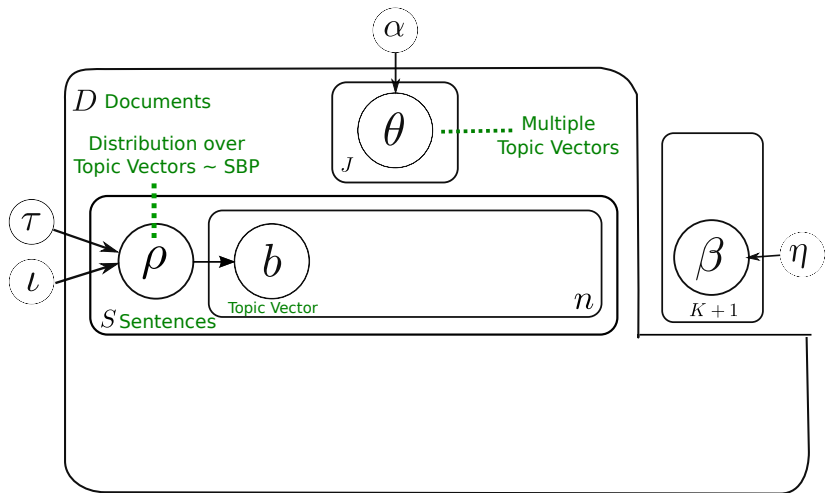
Specific Correspondence Topic Model (SCTM)



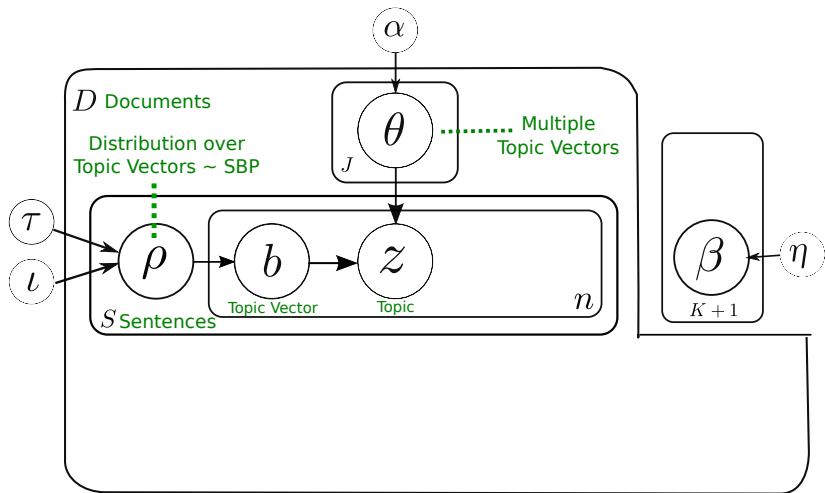
Specific Correspondence Topic Model (SCTM)



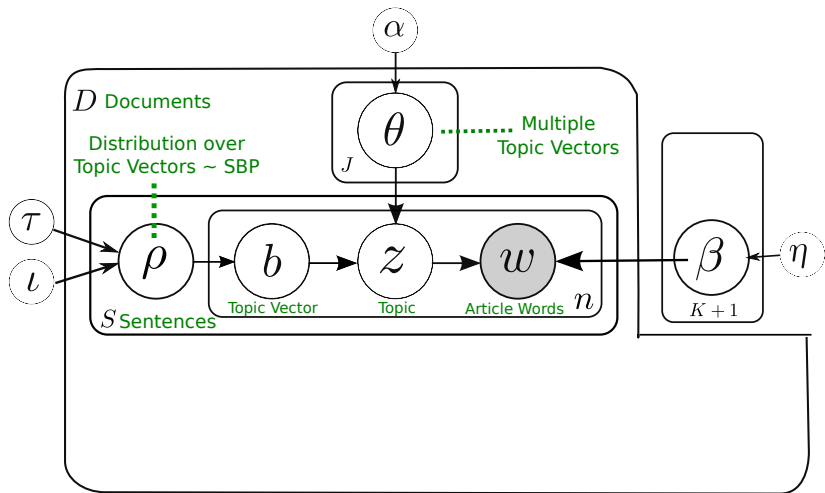
Specific Correspondence Topic Model (SCTM)



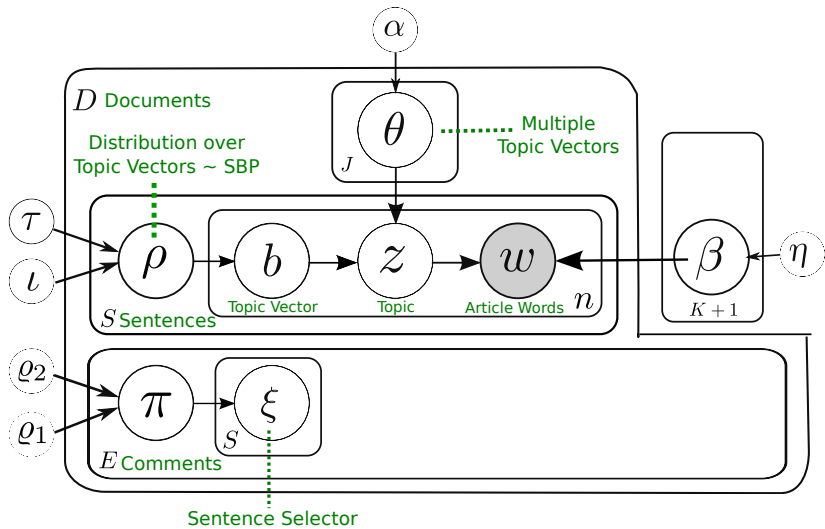
Specific Correspondence Topic Model (SCTM)



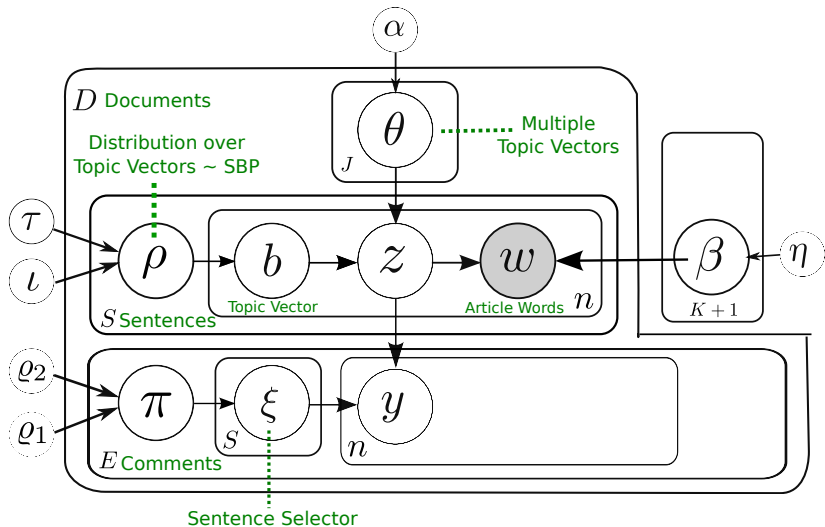
Specific Correspondence Topic Model (SCTM)



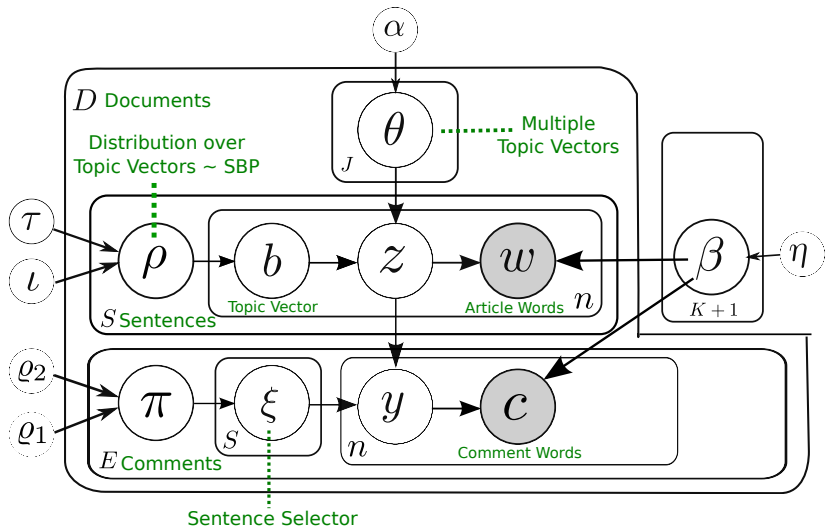
Specific Correspondence Topic Model (SCTM)



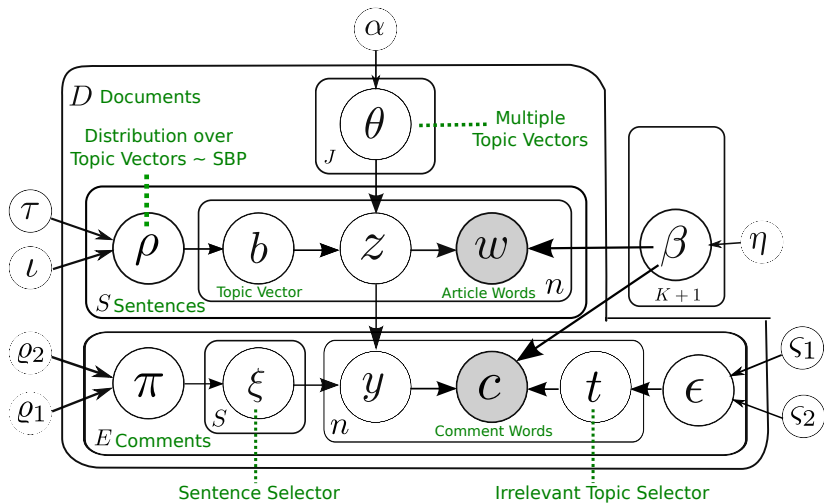
Specific Correspondence Topic Model (SCTM)



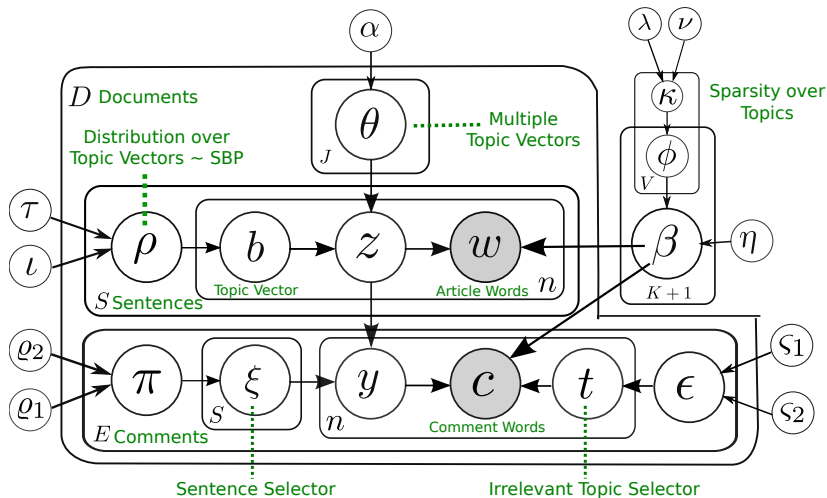
Specific Correspondence Topic Model (SCTM)



Specific Correspondence Topic Model (SCTM)



Specific Correspondence Topic Model (SCTM)



SBP is a suitable prior over MTV

Dirichlet is not appropriate

Over-estimation – general comments as specific

Under-estimation – specific comments as general

SBP is a suitable prior over MTV

Dirichlet is not appropriate

Over-estimation – general comments as specific

Under-estimation – specific comments as general

Stick-breaking process [5]

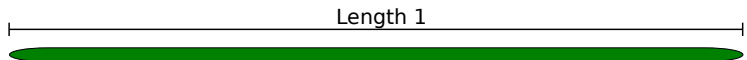
A random probability measure G is a stick-breaking prior if:

$G = \sum_{j=1}^J \rho_j \delta_{\theta_j}$, $\theta_j \stackrel{iid}{\sim} \Gamma$ and (ρ_j) are constructed as

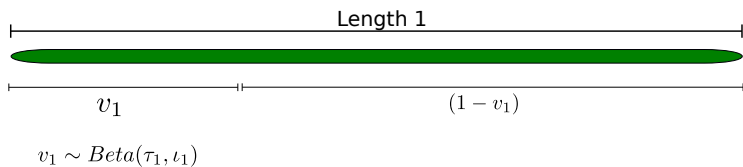
$\rho_1 = v_1$, $\rho_j = v_j \prod_{l=1}^{j-1} (1 - v_l)$, $j > 1$

where $v_j \stackrel{iid}{\sim} \text{Beta}(\tau_j, \nu_j)$, $j < J$, and $v_J = 1$

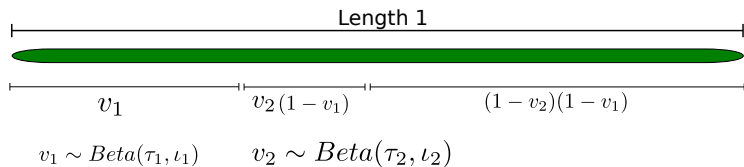
SBP – breaking a stick of unit length



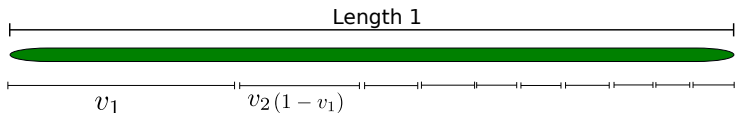
SBP – breaking a stick of unit length



SBP – breaking a stick of unit length



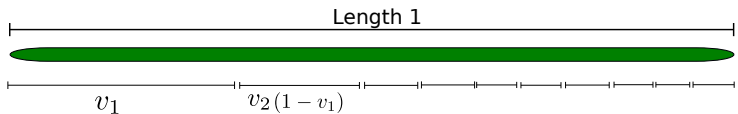
SBP – breaking a stick of unit length



$$v_1 \sim \text{Beta}(\tau_1, \iota_1) \quad v_2 \sim \text{Beta}(\tau_2, \iota_2) \quad v_i \sim \text{Beta}(\tau_i, \iota_i)$$

$$G = \sum_{i=1}^J \rho_i \delta_{\theta_i} \quad \rho_i = v_i \prod_{j=1}^{i-1} (1 - v_j) \quad \rho_J = 1 - \sum_{i=1}^{J-1} \rho_i$$

SBP – breaking a stick of unit length



$$v_1 \sim \text{Beta}(\tau_1, \iota_1) \quad v_2 \sim \text{Beta}(\tau_2, \iota_2) \quad v_i \sim \text{Beta}(\tau_i, \iota_i)$$

$$G = \sum_{i=1}^J \rho_i \delta_{\theta_i} \quad \rho_i = v_i \prod_{j=1}^{i-1} (1 - v_j) \quad \rho_J = 1 - \sum_{i=1}^{J-1} \rho_i$$

allows a large J , but higher indexed vectors are discouraged

Efficient easy to implement inference

Collapsed Gibbs sampling

Marginalize out the real valued variables

Challenge: Marginalizing SBP

Used the relationship between Generalized Dirichlet and SBP

GDD is conjugate to the multinomial distribution

Real life news & blogs for evaluation

3 datasets: two popular blogs and one popular news site

Blog 1: ArsTechnica Science (arstechnica.com/science)

1,369 articles & 90,654 comments (June 2011 to March 2013)

Blog 2: ArsTechnica Gadgets (arstechnica.com/gadgets)

2,186 articles & 160,761 comments (August 2011 to August 2013)

News: Yahoo! news (news.yahoo.com)

730 articles & 138,538 comments
(most-commented articles before 31 March 2013)

Gold-standard (manually annotated)

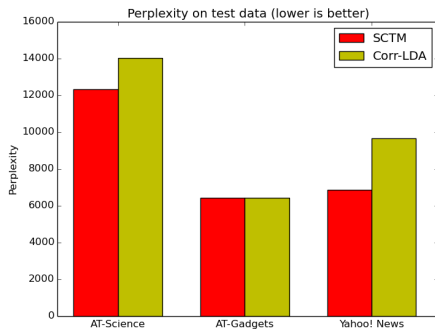
500 articles with 3176 comments from ArsTechnica Science
(March 2012 to March 2013)

Resources available at mllab.csa.iisc.ernet.in/sctm

SCTM outperforms in perplexity – popular metric

- **Perplexity** of a test data D_{test} of M articles is defined as:

$$perp(D_{test}) = \exp\left(-\frac{\sum_{d=1}^M \log p(w_d)}{\sum_{d=1}^M N_d}\right)$$

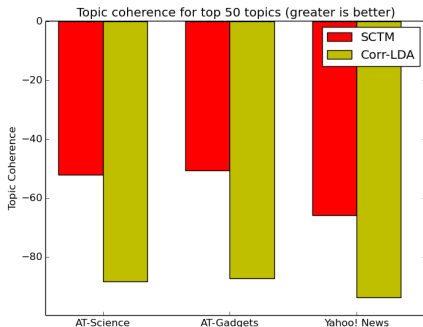


SCTM is better in topic quality

- **Topic Coherence** on top σ words

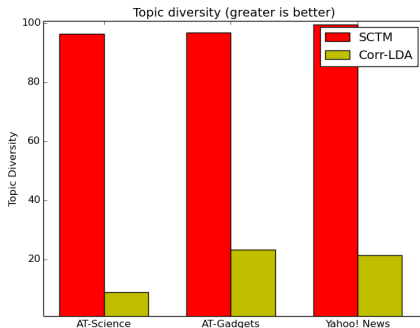
$$TC(\sigma) = \sum_{i \leq \sigma} \sum_{j < i} \log \frac{D(w_i, w_j) + \epsilon}{D(w_j)}$$

$D(w)$: document frequency of word w ; $D(w_i, w_j)$: document frequency of w_i and w_j taken together



SCTM discovers a widely diverse set of topics

- **Topic Diversity** is average textual difference between topic i and topic j $Tdiv_{ij} = 1 - \frac{|tw_i \cap tw_j|}{5}$
 tw_i : set of top 5 words for topic i



SCTM is substantially superior in SCL

Table: Detecting specific comments

Model	Precision	Recall	F1
SCTM	0.60	0.61	0.60
Corr-LDA	0.23	0.21	0.22

Table: Aligning specific comments to the sentences

Model	Precision	Recall	F1
SCTM	0.365	0.442	0.400
Corr-LDA	0.019	0.038	0.025

Using SCTM for listening to the user

Analyze what people say

Article

(sentences are color coded by topic, red-box is hot-spot)

Obama heads to Middle East with low expectations

When President Barack Obama steps into the Middle East's political cauldron this coming week, he won't be seeking any grand resolution for the region's vexing problems. His goal will be trying to keep the troubles, from Iran's suspected pursuit of a nuclear weapon to the bitter discord between Israelis and Palestinians, from boiling over on his watch.

This is not about accomplishing anything now. This is what I call a down payment trip," said Aaron David Miller, an adviser on Mideast peace to six secretaries of state

For much of Obama's first term, White House officials saw little reason for him to go to the region without a realistic chance for a peace accord between the Israelis and Palestinians.....

Officials now see the lowered expectations as a chance to create space for frank conversations between Obama and both sides about what it will take to get back to the negotiating table.

Netanyahu, in a speech to the United Nations in September, said Iran was about six months away from being able to build a bomb. Obama told an Israeli television station this past week that the U.S. thinks it would take "over a year or so for Iran to actually develop a nuclear weapon."

Traveling to the West Bank, Obama will meet with Palestinian Authority President Mahmoud Abbas and Prime Minister Salam Fayyad in Ramallah. Obama and Fayyad will visit a Palestinian youth center, another attempt to reach the region's young people..

Obama will make a 24-hour stop in Jordan, an important U.S. ally, where the president's focus will be on the violence in neighboring Syria. More than 450,000 Syrians have fled to Jordan, crowding refugee camps and overwhelming aid organizations.

Comments

(color coded with parts of article and topic)

I am not sure why he is going. To go under the gloom of 'low expectations' indicates a wasted trip. Though Obama seems more comfortable in some Middle Eastern countries than in US States. Sure spend more money we do not have

Who is John Galt?

THEY have vexing problems.

If you want to solve the Palestinian problem it's simple, create jobs and security for the people. When a man has a job and family to provide for, he doesn't think of war. He wants to live and let live. The Palestinian leadership are getting rich off the conflict between Israel and Palestine! Arafat died with a BILLION DOLLARS IN THE BANK!!! I wonder if Netanyahu plays golf?

This problem with Iran's nuclear weapons program should have been handled long before the tipping point of Crisis.....Obama waited too long.

SCTM Topics (color coded with article and comments)

Obama Middle East Visit
israel-Palestine Peace Talks
Iran-Israel Nuclear Dispute
Obama Visit Expectations
Syria & Arab Spring

obama, east, america, israeli, netanyahu, trip, americans, american, visit, country
israel, peace, palestinians, land, palestinian, muslim, netanyahu, zionist, world, hate, israelis
iran, israel, nuclear, world, military, iranian, peace, weapons, destroyer, attack
expectations, middle, obama, money, president, east, office, barry, hope, waste, biden, troubles
muslim, stay, middle, muslims, syrians, lebanon, east, egypt, arab, israel, obama, sunni

Corr-LDA Topic (mixed topic)

Obama Visit & Mideast Issues
israel, obama, jews, peace, palestinians, jewish, israeli, land, president, east

- News reporting President Obama's visit to the Middle East during March 19-23, 2013
- At the time of crawling, the article had 1919 comments (most-commented)

Using SCTM for listening to the user

Analyze what people say

Article

(sentences are color coded by topic, red-box is hot-spot)

Obama heads to Middle East with low expectations

When President Barack Obama steps into the Middle East's political cauldron this coming week, he won't be seeking any grand resolution for the region's vexing problems. His goal will be trying to keep the troubles, from Iran's suspected pursuit of a nuclear weapon to the bitter discord between Israelis and Palestinians, from boiling over on his watch.

This is not about accomplishing anything now. This is what I call a down payment trip," said Aaron David Miller, an adviser on Mideast peace to six secretaries of state

For much of Obama's first term, White House officials saw little reason for him to go to the region without a realistic chance for a peace accord between the Israelis and Palestinians.....

Officials now see the lowered expectations as a chance to create space for frank conversations between Obama and both sides about what it will take to get back to the negotiating table.

Netanyahu, in a speech to the United Nations in September, said Iran was about six months away from being able to build a bomb. Obama told an Israeli television station this past week that the U.S. thinks it would take "over a year or so for Iran to actually develop a nuclear weapon."

Traveling to the West Bank, Obama will meet with Palestinian Authority President Mahmoud Abbas and Prime Minister Salam Fayyad in Ramallah. Obama and Fayyad will visit a Palestinian youth center, another attempt to reach the region's young people..

Obama will make a 24-hour stop in Jordan, an important U.S. ally, where the president's focus will be on the violence in neighboring Syria. More than 450,000 Syrians have fled to Jordan, crowding refugee camps and overwhelming aid organizations.

Comments

(color coded with parts of article and topic)

I am not sure why he is going. To go under the gloom of 'low expectations' indicates a wasted trip. Though Obama seems more comfortable in some Middle Eastern countries than in US States. Sure spend more money we do not have

Who is John Galt?

THEY have vexing problems.

If you want to solve the Palestinian problem it's simple, create jobs and security for the people. When a man has a job and family to provide for, he doesn't think of war. He wants to live and let live. The Palestinian leadership are getting rich off the conflict between Israel and Palestine! Arafat died with a BILLION DOLLARS IN THE BANK!!

I wonder if Netanyahu plays golf?

This problem with Iran's nuclear weapons program should have been handled long before the tipping point of Crisis.....Obama waited too long.

SCTM Topics (color coded with article and comments)

Obama Middle East Visit	obama, east, america, israeli, netanyahu, trip, americans, american, visit, country
Israel-Palestine Peace Talks	israel, peace, palestinians, land, palestinian, muslim, netanyahu, zionist, world, hate, israelis
Iran-Israel Nuclear Dispute	iran, israel, nuclear, world, military, iranian, peace, weapons, destroyer, attack
Obama Visit Expectations	expectations, middle, obama, money, president, east, office, barry, hope, waste, biden, troubles
Syria & Arab Spring	muslim, stay, middle, muslims, syrians, lebanon, east, egypt, arab, israel, obama, sunni

Corr-LDA Topic (mixed topic)

Obama Visit & Mideast Issues	israel, obama, jews, peace, palestinians, jewish, israeli, land, president, east
------------------------------	--

- We discover fine topics in article like “Obama’s MidEast Visit”, “**Visit Expectation**”, “**Israel-Palestine Peace**” and “**Iran Nuclear Dispute**”
- Among the specific comments, the most discussed topic is “**Visit Expectation**” (red colored) with **195 comments**

Using SCTM for listening to the user

Analyze what people say

Article

(sentences are color coded by topic, red-box is hot-spot)

Obama heads to Middle East with low expectations

When President Barack Obama steps into the Middle East's political cauldron this coming week, he won't be seeking any grand resolution for the region's vexing problems. His goal will be trying to keep the troubles, from Iran's suspected pursuit of a nuclear weapon to the bitter discord between Israelis and Palestinians, from boiling over on his watch.

"This is not about accomplishing anything now. This is what I call a down payment trip," said Aaron David Miller, an adviser on Mideast peace to six secretaries of state

For much of Obama's first term, White House officials saw little reason for him to go to the region without a realistic chance for a peace accord between the Israelis and Palestinians.....

Officials now see the lowered expectations as a chance to create space for frank conversations between Obama and both sides about what it will take to get back to the negotiating table.

Netanyahu, in a speech to the United Nations in September, said Iran was about six months away from being able to build a bomb. Obama told an Israeli television station this past week that the U.S. thinks it would take "over a year or so for Iran to actually develop a nuclear weapon."

Traveling to the West Bank, Obama will meet with Palestinian Authority President Mahmoud Abbas and Prime Minister Salam Fayyad in Ramallah. Obama and Fayyad will visit a Palestinian youth center, another attempt to reach the region's young people..

Obama will make a 24-hour stop in Jordan, an important U.S. ally, where the president's focus will be on the violence in neighboring Syria. More than 450,000 Syrians have fled to Jordan, crowding refugee camps and overwhelming aid organizations.

Comments

(color coded with parts of article and topic)

I am not sure why he is going. To go under the bloom of 'low expectations' indicates a wasted trip. Though Obama seems more comfortable in some Middle Eastern countries than in US States. Sure spend more money we do not have

Who is John Galt?

THEY have vexing problems.

If you want to solve the Palestinian problem it's simple, create jobs and security for the people. When a man has a job and family to provide for, he doesn't think of war. He wants to live and let live. The Palestinian leadership are getting rich off the conflict between Israel and Palestine! Arafat died with a BILLION DOLLARS IN THE BANK!!!

I wonder if Netanyahu plays golf?

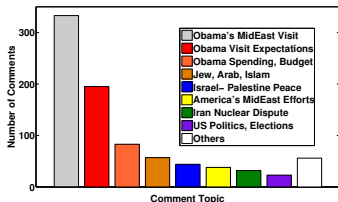
This problem with Iran's nuclear weapons program should have been handled long before the tipping point of Crisis.....Obama waited too long.

SCTM Topics (color coded with article and comments)

Obama Middle East Visit	obama, east, america, israeli, netanyahu, trip, americans, american, visit, country
Israel-Palestine Peace Talks	israel, peace, palestinians, land, palestinian, muslim, netanyahu, zionist, world, hate, israelis
Iran-Israel Nuclear Dispute	iran, israel, nuclear, world, military, iranian, peace, weapons, destroyer, attack
Obama Visit Expectations	expectations, middle, obama, money, president, east, office, barry, hope, waste, biden, troubles
Syria & Arab Spring	muslim, stay, middle, muslims, syrians, lebanon, east, egypt, arab, israel, obama, sunni

Corr-LDA Topic (mixed topic)

Obama Visit & Mideast Issues	israel, obama, jews, peace, palestinians, jewish, israel, land, president, east
------------------------------	---



Generate comment statistics for an article based on topic

Using SCTM for listening to the user

Read comments along with the articles

Article

(sentences are color coded by topic, red-box is hot-spot)

Obama heads to Middle East with low expectations

When President Barack Obama steps into the Middle East's political cauldron this coming week, he won't be seeking any grand resolution for the region's vexing problems. His goal will be trying to keep the troubles, from Iran's suspected pursuit of a nuclear weapon to the bitter discord between Israelis and Palestinians, from boiling over on his watch.

This is not about accomplishing anything now. This is what I call a down payment trip," said Aaron David Miller, an adviser on Mideast peace to six secretaries of state

For much of Obama's first term, White House officials saw little reason for him to go to the region without a realistic chance for a peace accord between the Israelis and Palestinians.....

Officials now see the lowered expectations as a chance to create space for frank conversations between Obama and both sides about what it will take to get back to the negotiating table.

Netanyahu, in a speech to the United Nations in September, said Iran was about six months away from being able to build a bomb. Obama told an Israeli television station this past week that the U.S. thinks it would take "over a year or so for Iran to actually develop a nuclear weapon."

Traveling to the West Bank, Obama will meet with Palestinian Authority President Mahmoud Abbas and Prime Minister Salam Fayyad in Ramallah. Obama and Fayyad will visit a Palestinian youth center, another attempt to reach the region's young people.....

Obama will make a 24-hour stop in Jordan, an important U.S. ally, where the president's focus will be on the violence in neighboring Syria. More than 450,000 Syrians have fled to Jordan, crowding refugee camps and overwhelming aid organizations.

Comments

(color coded with parts of article and topic)

I am not sure why he is going. To go under the gloom of 'low expectations' indicates a wasted trip. Though Obama seems more comfortable in some Middle Eastern countries than in US States. Sure spend more money we do not have

Who is John Galt?

THEY have vexing problems.

If you want to solve the Palestinian problem it's simple, create jobs and security for the people. When a man has a job and family to provide for, he doesn't think of war. He wants to live and let live. The Palestinian leadership are getting rich off the conflict between Isreal and Palestine! Arafat died with a BILLION DOLLARS IN THE BANK!!

I wonder if Netanyahu plays golf?

This problem with Iran's nuclear weapons program should have been handled long before the tipping point of Crisis.....Obama waited too long.

SCTM Topics (color coded with article and comments)

Obama Middle East Visit
israel, peace, palestinians, land, palestinian, muslim, netanyahu, zionist, world, hate, israelis
Iran-Israel Nuclear Dispute
iran, israel, nuclear, world, military, iranian, peace, weapons, destroyer, attack
Obama Visit Expectations
expectations, middle, obama, money, president, east, office, barry, hope, waste, biden, troubles
Syria & Arab Spring
muslim, stay, middle, muslims, syrians, lebanon, east, egypt, arab, israel, obama, sunni

Corr-LDA Topic (mixed topic)

Obama Visit & Mideast Issues
israel, obama, jews, peace, palestinians, jewish, israeli, land, president, east

red colored comment can be popped up while reading red colored text

Using SCTM for listening to the user

Discover hot-spots in the articles

Article

(sentences are color coded by topic, red-box is hot-spot)

Obama heads to Middle East with low expectations

When President Barack Obama steps into the Middle East's political cauldron this coming week, he won't be seeking any grand resolution for the region's vexing problems. His goal will be trying to keep the troubles, from Iran's suspected pursuit of a nuclear weapon to the bitter discord between Israelis and Palestinians, from boiling over on his watch.

This is not about accomplishing anything now. This is what I call a down payment trip," said Aaron David Miller, an adviser on Mideast peace to six secretaries of state

For much of Obama's first term, White House officials saw little reason for him to go to the region without a realistic chance for a peace accord between the Israelis and Palestinians.....

Officials now see the lowered expectations as a chance to create space for frank conversations between Obama and both sides about what it will take to get back to the negotiating table.

Netanyahu, in a speech to the United Nations in September, said Iran was about six months away from being able to build a bomb. Obama told an Israeli television station this past week that the U.S. thinks it would take "over a year or so for Iran to actually develop a nuclear weapon."

Traveling to the West Bank, Obama will meet with Palestinian Authority President Mahmoud Abbas and Prime Minister Salam Fayyad in Ramallah. Obama and Fayyad will visit a Palestinian youth center, another attempt to reach the region's young people..

Obama will make a 24-hour stop in Jordan, an important U.S. ally, where the president's focus will be on the violence in neighboring Syria. More than 450,000 Syrians have fled to Jordan, crowding refugee camps and overwhelming aid organizations.

Comments

(color coded with parts of article and topic)

I am not sure why he is going. To go under the gloom of 'low expectations' indicates a wasted trip. Though Obama seems more comfortable in some Middle Eastern countries than in US States. Sure spend more money we do not have

Who is John Galt?

THEY have vexing problems.

If you want to solve the Palestinian problem it's simple, create jobs and security for the people. When a man has a job and family to provide for, he doesn't think of war. He wants to live and let live. The Palestinian leadership are getting rich off the conflict between Israel and Palestine! Arafat died with a BILLION DOLLARS IN THE BANK!!!

I wonder if Netanyahu plays golf?

This problem with Iran's nuclear weapons program should have been handled long before the tipping point of Crisis.....Obama waited too long.

SCTM Topics (color coded with article and comments)

Obama Middle East Visit	obama, east, america, israeli, netanyahu, trip, americans, american, visit, country
Israel-Palestine Peace Talks	israel, peace, palestinians, land, palestinian, muslim, netanyahu, zionist, world, hate, israelis
Iran-Israel Nuclear Dispute	iran, israel, nuclear, world, military, iranian, peace, weapons, destroyer, attack
Obama Visit Expectations	expectations, middle, obama, money, president, east, office, barry, hope, waste, biden, troubles
Syria & Arab Spring	muslim, stay, middle, muslims, syrians, lebanon, east, egypt, arab, israel, obama, sunni

Corr-LDA Topic (mixed topic)

Obama Visit & Mideast Issues	israel, obama, jews, peace, palestinians, jewish, israeli, land, president, east
------------------------------	--

Hot-Spot: most commented part of the article (red colored box)

Using SCTM for listening to the user

Comment cleaning

Filter out irrelevant comments for better user experience

Voyager over the "heliocliff," but Solar System transition mysterious

- *You sure? Maybe they were just clicking that "restart later" button every four hours...*
- *If I ever have kids, this is what they're growing up to. Old McDonald had a cliff, e-i-e-i-o And on that cliff there was a Voyager, e-i-e-i-o*

Using SCTM for listening to the user

Recommending articles to comment

Recommend articles for commenting based on comment topic

Article

Obama heads to Middle East with low expectations

When President Barack Obama steps into the Middle East's political cauldron this coming week, he won't be seeking any grand resolution for the region's vexing problems

.....
Netanyahu, in a speech to the United Nations in September, said Iran was about six months away from being able to build a bomb. Obama told an Israeli television station this past week that the U.S. thinks it would take "over a year or so for Iran to actually develop a nuclear weapon."
.....

Comments



User XYZ

This problem with Iran's nuclear weapons program should have been handled long before the tipping point of Crisis.....Obama waited too long.

→ **XYZ, your comment has been received.**

You may also like to comment on the following articles:

"Iran launches destroyer in the Caspian Sea"

"Obama tells Iranians now is the time to settle nuclear dispute"

"Can Iran, world powers build on recent progress in nuclear talks?"

"Iran: Commanders authorized to respond to attacks"

"Obama won't trip over Netanyahu's Iran "red line"

References I

- [1] David M. Blei and Michael I. Jordan.
Modeling annotated data.
In *SIGIR*, pages 127–134. ACM, 2003.
- [2] David M. Blei, Andrew Y. Ng, and Michael I. Jordan.
Latent Dirichlet allocation.
Journal of Machine Learning Research, 3:993–1022,
January 2003.
- [3] Mrinal Das, Suparna Bhattacharya, Chiranjib
Bhattacharyya, and Gopinath Kanchi.
Subtle topic models and discovering subtly manifested
software concerns automatically.
In *Proceedings of the 30th International Conference on
Machine Learning (ICML-13)*, pages 253–261, 2013.

- [4] Yuheng Hu, Ajita John, Dorée Duncan Seligmann, and Fei Wang.
What were the tweets about? Topical associations between public events and twitter feeds.
In *ICWSM*, 2012.
- [5] Hemant Ishwaran and Lancelot F James.
Gibbs sampling methods for stick-breaking priors.
Journal of the American Statistical Association,
96(453):161–173, 2001.
- [6] Zongyang Ma, Aixin Sun, Quan Yuan, and Gao Cong.
Topic-driven reader comments summarization.
In *CIKM*, pages 265–274. ACM, 2012.