



# Extracting Problem and Resolution Information from Online Discussion Forums

Preethi Raghavan\*, Rose Catherine, Shajith Ikbal, Nanda Kambhatla, Debapriyo Majumdar

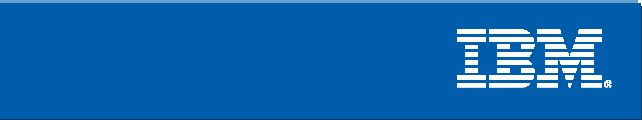
IBM Research

(\* - Ohio State University, USA)



## Introduction

- ❖ Social media and user-generated content have grown rapidly over the past few years
  - Blogs, social networking sites, online discussion forums
- ❖ A large fraction of these are in the form of question - answer discussions, where many members of the forum collaborate to solve a problem.
- ❖ They provide an excellent source of information to users seeking solutions to same or similar problems



hotel near airport bangalore ✕ Search

About 26,400 results (0.42 seconds) [Advanced search](#)

[Bangalore: sleep near airport? - India Travel Forum | IndiaMike.com](#)

has anyone heard of Royal Inn **Hotel near airport** or has anyone used .... **Hotel near Bangalore airport**, jodiesner, **Bangalore**, 1, Jan 17th, ...  
[www.indiamike.com](#) > ... > [Cities in India](#) > [Bangalore](#) - [Cached](#) - [Similar](#)

[Hotels Near Bangalore International Airport - Hotels in India ...](#)

9 posts - 7 authors - Last post: 14 Sep  
 The Basil Ikon **Hotel** and The Leela Palace Kempinski **Bangalore** are located at a distance of 3 km from **Bangalore International Airport**. ...  
[www.sukh-duk.com/forums/showthread.php?t=10368](#) - [Cached](#)

[plots near bangalore international airport - Click India Classifieds](#)

10 posts - 1 author - Last post: 23 Nov  
 at Daily breakfast at Round trip Transfers **Airport-Hotel-Airport** ( SIC) ...  
 MEASURING 3600sq.ft **NEAR AIRPORT ROAD**. • 13-Sep-2010 • **Bangalore** ...  
[www.clickindia.com/search.php?q...near+bangalore...airport](#) - [Cached](#) - [Similar](#)

[BLR hotel near airport? - FlyerTalk Forums](#)

13 posts - 7 authors - Last post: 16 Nov  
 I'm flying to **Bangalore** arriving very early am (00.25). ... Are there any decent hotels due to open **near** the **BLR airport**? ...  
[www.flyertalk.com/forum/india/1002266-blr-hotel-near-airport.html](#) - [Cached](#)

[Bangalore Airport: Hangout or Hotel? - Lonely Planet travel forum](#)

11 posts - 5 authors - Last post: 20 Jan 2008  
 My husband and I arrive in **Bangalore** from Thailand around 3am. ... I dont quite know of any **hotel near** the **airport** which falls into that ...  
[www.lonelyplanet.com/thomtree/thread.jspta?threadID=1475234](#) - [Cached](#)

[resorts near bangalore - Click India Classifieds](#)

10 posts - 5 authors - Last post: 9 Nov  
 Hotels in India, Business Hotels **near Bangalore International Airport**. • 16-Sep-2010 • **Bangalore**. Solitaire **Hotel** located in **near Bangalore** ...  
[www.clickindia.com/search.php?q=resorts+near+bangalore](#) - [Cached](#) - [Similar](#)



hotel near airport bangalore

X Search

About 26,400 results (0.42 seconds)

Advanced search

[Bangalore: sleep near airport? - India Travel Forum | IndiaMike.com](#)

has anyone heard of Royal Inn **Hotel near airport** or has anyone used .... **Hotel near Bangalore airport**, jodiesner, **Bangalore**, 1, Jan 17th, ...  
[www.indiamike.com](#) > ... > Cities in India > Bangalore - [Cached](#) - [Similar](#)

[Hotels Near Bangalore International Airport - Hotels in India ...](#)

9 posts - 7 authors - Last post: 14 Sep  
The Basil Ikon **Hotel** and The Leela Palace Kempinski **Bangalore** are located at a distance of 3 km from **Bangalore International Airport**. ...  
[www.sukh-duk.com/forums/showthread.php?t=10368](#) - [Cached](#)

[plots near bangalore international airport - Click India Classifieds](#)

10 posts - 1 author - Last post: 23 Nov  
at Daily breakfast at Round trip Transfers **Airport-Hotel-Airport** ( SIC) ...  
MEASURING 3600sq.ft **NEAR AIRPORT ROAD**. • 13-Sep-2010 • **Bangalore** ...  
[www.clickindia.com/search.php?q...near+bangalore...airport](#) - [Cached](#) - [Similar](#)

[BLR hotel near airport? - FlyerTalk Forums](#)

13 posts - 7 authors - Last post: 16 Nov  
I'm flying to **Bangalore** arriving very early am (00.25). ... Are there any decent hotels due to open **near** the **BLR airport**? ...  
[www.flyertalk.com/forum/india/1002266-blr-hotel-near-airport.html](#) - [Cached](#)

[Bangalore Airport: Hangout or Hotel? - Lonely Planet travel forum](#)

11 posts - 5 authors - Last post: 20 Jan 2008  
My husband and I arrive in **Bangalore** from Thailand around 3am. ... I dont quite know of any **hotel near** the **airport** which falls into that ...  
[www.lonelyplanet.com/thorntree/thread.jspa?threadID=1475234](#) - [Cached](#)

[resorts near bangalore - Click India Classifieds](#)

10 posts - 5 authors - Last post: 9 Nov  
Hotels in India, Business Hotels **near Bangalore International Airport**. • 16-Sep-2010 • **Bangalore**. Solitaire **Hotel** located in **near Bangalore** ...  
[www.clickindia.com/search.php?q=resorts+near+bangalore](#) - [Cached](#) - [Similar](#)



## The Leela Palace

[www.sukh-duk.com/forums/showthread.php?t=10368](#)

## The Oberoi

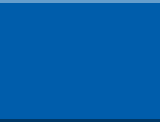
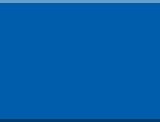
[www.clickindia.com/search.php?q...near+bangalore...airport](#)

## The Park Hotel

[www.lonelyplanet.com/thorntree/thread.jspa?threadID=1475234](#)

## Royal Orchid Hotel

[www.indiamike.com](#) > ... > Cities in India > Bangalore



scp without password in linux



Search

Instant is

About 28,900 results (0.36 seconds)

Advanced search

▶ [SCP from one server to another without password prompt - Stack ...](#) 🔍

Solved - 1 answer

Browse other questions tagged **linux passwords scp** or ask your own question. ... How can I transfer a file using **scp without a password** from a Perl script? ...

[stackoverflow.com/.../scp-from-one-server-to-another-without-password-prompt](#) -

Cached - Similar

[scp with password - HowtoForge Forums | HowtoForge - Linux Howtos ...](#) 🔍

3 posts - 3 authors - Last post: 5 Mar 2007

**scp with password** Technical. ... English | HowtoForge.de - **Linux-Howtos auf Deutsch** ... and there you go: no more **password** typing. ...

[www.howtoforge.com > ... > Linux Forums > Technical](#) - Cached - Similar

[ssh/scp without password not working](#) 🔍

Hello all, I'm having problems getting **ssh/scp** working **without** prompting for a **password**.

What I did is the following: \*) installed openssh \*)

[www.linuxquestions.org/.../linux.../ssh-scp-without-password-not-working-307388/](#) - Cached

[sftp/scp without password - Linux Forums](#) 🔍

2 posts - 2 authors - Last post: 19 Aug 2009

How can I use **sftp/scp without password in linux**. How can I use sftp/scp to write script and execute in cronjob ?? any example or sample ...

[www.linuxforums.org/...linux/151687-sftp-scp-without-password.html](#) - Cached

[step by step SSH/SCP without password - maemo.org - Talk](#) 🔍

8 posts - 5 authors - Last post: 12 Jan 2009

Step by step **SSH/scp Linux** and your tablet, **without password**. I know NOTHING about security, you got it right. You may open your tablet to ...

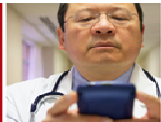
[talk.maemo.org/showthread.php?t=26157](#) - Cached

[OTN Discussion Forums : set up scp without password ...](#) 🔍

5 posts - 2 authors - Last post: 11 Aug

we have a requirement to **scp** file to a **LINUX** machine from our Solaris 9 ... Re: set up **scp without password**. Posted: Aug 11, 2010 5:50 PM in ...

[forums.oracle.com/forums/thread.jspa?messageID=8161304](#) - Cached



[sunForumsGuest19](#)

Posts: 37,771  
Registered: 10/22/10

### set up scp without password

Posted: Jan 22, 2009 7:31 PM



Hi,  
we have a requirement to scp file to a LINUX machine from our Solaris 9 server. So far, I did

```
$ ssh-keygen -t dsa
Enter file in which to save the key(/export/home/user/.ssh/id_rsa):
Generating public/private rsa key pair.
Enter passphrase(empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /export/home/user/.ssh/id_rsa.
Your public key has been saved in /export/home/user/.ssh/id_rsa.pub.
The key fingerprint is:
md5 1024 56:0b:aa:4a:43:ee:e4:79:54:a3:2a:5d:06:20:d0:8d user@server
```

```
$ ls -ltrh .ssh/
total 14
-rwx----- 1 user oinstall 235 Feb 28 2007 authorized_keys
-rwx----- 1 user oinstall 1.7K Jan 22 17:17 known_hosts
-rw-r--r-- 1 user oinstall 224 Jan 23 12:55 id_rsa.pub
-rw----- 1 user oinstall 887 Jan 23 12:55 id_rsa
-rw-r--r-- 1 user oinstall 604 Jan 23 13:36 id_dsa.pub
-rw----- 1 user oinstall 668 Jan 23 13:36 id_dsa
```

```
$ ssh-copy-id -i ~/.ssh/id_dsa.pub user@labt88fms
ksh: ssh-copy-id: not found
```

```
$ ssh user@labt88fms
user@labt88fms password:
```

I am still being asked for a password. Once I enter the pwd I do ls on the remote box and see

```
id_rsa.pub
```

Can someone please advise how can I get the scp to work so that I do not get asked for a password via shell script. thank you

[user4994457](#)

### Re: set up scp without password



[sunForumsGuest19](#)   
 Posts: 37,771  
 Registered: 10/22/10

**set up scp without password**   [Reply](#)  
 Posted: Jan 22, 2009 7:31 PM

Hi,  
 we have a requirement to scp file to a LINUX machine from our Solaris 9 server. So far, I did



```
$ ssh-keygen -t dsa
Enter file in which to save the key(/export/home/user/.ssh/id_rsa):
Generating public/private rsa key pair.
Enter passphrase(empty for no passphrase):
```

Please read through the SSH documentation, especially the sshd man page and the section on Authorized Keys.

The public key can't just be sitting in the home directory. The contents have to be inserted into a file. The normal location of the file is ~/.ssh/authorized\_keys, but that could be changed by the sshd\_config file.



--  
 Darren

[sunForumsGuest19](#)   
 Posts: 37,771  
 Registered: 10/22/10

**Re: set up scp without password**   [Reply](#)  
 Posted: Aug 11, 2010 2:00 PM  
 ↑ in response to: [user4994457](#)

Note that ssh-copy-id failed. Solaris doesn't appear to have Linux's ssh-copy-id command. use ftp, rsync or scp to deliver the file.

[sunForumsGuest19](#)   
 Posts: 37,771  
 Registered: 10/22/10


**Re: set up scp without password**   [Reply](#)  
 Posted: Aug 11, 2010 5:45 PM  
 ↑ in response to: [sunForumsGuest19](#)

Most likely the permission on the home directory.

They most likely need to be

```
chmod 755
```

Note, if you allow group or other writes, then other people can modify your SSH keys.

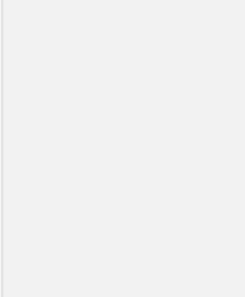
[user4994457](#) 

[sunForumsGuest19](#)   
 Posts: 37,771  
 Registered: 10/22/10

**set up s**  
 Posted: .

Hi,  
 we have  
 far, I did

\$ ssh-key  
 Enter file  
 Generati  
 Enter pas




[sunForumsGuest19](#)   
 Posts: 37,771  
 Registered: 10/22/10

[sunForumsGuest19](#)   
 Posts: 37,771  
 Registered: 10/22/10

[user4994457](#) 

Thread Tools ▾ Display Modes ▾

5th March 2007, 02:19 #1

[knowram](#)  Member  
 Join Date: Feb 2007  
 Posts: 33  
 Thanks: 0  
 Thanked 1 Time in 1 Post


 **scp with password**


Is there anyway to add a password into the scp command string?  
 something like.

```
scp file me@192.168.0.101 password location
```



5th March 2007, 09:25 #2

[martinfst](#)  Senior Member  
 Join Date: Dec 2006  
 Location: Hilversum, The Netherlands  
 Posts: 880  
 Thanks: 1  
 Thanked 16 Times in 15 Posts

 If you don't want to enter password, you will need to use keys. Use ssh-keygen to create a private/public key pair, copy the public part to the destination server (\$HOME/.ssh/authorized\_keys) and there you go: no more password typing.

```
-----
```

Most likely the permission on the home directory.

They most likely need to be

```
chmod 755
```

Note, if you allow group or other writes, then other people can modify your SSH keys.



**sunForumsGuest19**  
 Posts: 37,771  
 Registered: 10/22/10

**set up s**  
 Posted: .  
 Hi,  
 we have  
 far, I did  
 \$ ssh-key  
 Enter file  
 Generatir  
 Enter pas

5th March 2007, 02:19  
**knowram**  
 Member  
**scp with password**  
 Is there anyway to add something like.  
 scp file me@192.168.0.1

5th March 2007, 09:25  
**martinfst**  
 Senior Member  
 If you don't want to ent key pair, copy the public no more password typing

**sunForumsGuest19**  
 Posts: 37,771  
 Registered: 10/22/10

**sunForumsGuest19**  
 Posts: 37,771  
 Registered: 10/22/10

**user4994457**

Most likely the permission c  
 They most likely need to be  
 chmod 755  
 Note, if you allow group or keys.

Author	Subject: scp without password
<a href="#">steven chang</a>	Feb 9, 2004 03:33:24 GMT hello : I am trying to scp from solaris 8 to redhat 8, and openssh version running on solaris 8 is OpenSSH_3.6.1p1 redhat 8 is OpenSSH_3.4p1, I generate the public/private key for solaris and ftp the id_dsa to redhat and use the id_dsa.pub to generate authority_keys file to redhat 8 . I already put ssh-agent start script in the .profile on the solaris 8 and auto give the passphrase through ssh-add, but I need update ssh-add pass phrase every time when I login the solaris. Now I want to write a script to scp a tar file from redhat 8. How can I auto give the passprase with interactive giving through the terminal input.That is how can writing a script to scp with giving any imformation and run this script in the crontab. Thank you very much steven chang

Note: If you are the author of this question and wish to assign points to any of the answers, please login first.For more information on assigning points ,click [here](#)

Sort Answers By: Date or [Points](#)

<a href="#">Mark Grant</a>	Feb 9, 2004 03:42:35 GMT 10 pts Have a look here, it might work for you. <a href="http://www.mail-archive.com/ssh%40clinet.fi/msg04511.html">http://www.mail-archive.com/ssh%40clinet.fi/msg04511.html</a>
<a href="#">steven chang</a>	Feb 9, 2004 03:52:13 GMT N/A: Question Author If I let the pass phrase empty when generating the public/private key , my goal would be achieved.



## Get relevant information / steps to solve a problem



1. Generate key-pair on the local-host using ssh-keygen
2. Install public key on the remote-host
3. Give appropriate permission to the .ssh directory on the remote-host



## Introduction – contd.

- ❖ Organizing and retrieving relevant information from such sources is not yet very effective
- ❖ Propose an approach to extract useful problem and resolution related information from discussion threads
- ❖ Study the effect of using different characteristics of the discussion, on the extraction process
- ❖ Results show considerable improvement over the baselines.



## Why do we care?

- ❖ Archived discussion threads provide a storehouse of useful information
- ❖ Information Retrieval:
  - ❑ Traditional keyword search tends to be inadequate as many distinguishing characteristics of discussion forums are ignored
    - Semi-structured
    - Span several pages
  - ❑ Identifying parts of the thread that have useful information aids in ranking relevant threads
- ❖ Effective presenting of information:
  - ❑ Exact resolution steps
    - Saves time
    - Applicability to self-help portals, technical support desks.



## Why is it difficult?

- ❖ Discussions are in natural language and mostly informal
- ❖ Many messages digress from the topic of discussion, adding noise to the text
- ❖ Problem and resolution text depends heavily on the domain
- ❖ Relevance is subjective



## General characteristics of a discussion thread

- ❖ Semi-structured unlike other web pages; conversational in nature.
- ❖ Discussion thread starts with a message with the problem description and some background information
- ❖ Followed by a series of replies with various suggestions towards resolving the problem.
- ❖ Flow of information is mostly sequential
  - ❑ Multiple responses to the original question or posts
  - ❑ Digressions where new questions are posed and answered
- ❖ Types of content:
  - ❑ Links to other threads, sites.
  - ❑ Problem author gives more information on the issue.
  - ❑ Feedback on whether a particular suggestion worked or not.
- ❖ May conclude with the problem author acknowledging successful resolution of the problem.



## Problem Statement

- ❖ Identify important parts of a discussion thread:
  - ❑ Problem related
  - ❑ Resolution related
  
- ❖ We cast this problem as a sequence labeling task:
  - ❑ Sentence level
  - ❑ Phrase level
  
- ❖ Train Conditional Random Fields (CRF) for the classification task
  - ❑ CRFs are probabilistic discriminative models for segmenting and labeling sequence data.
  - ❑ They predict the probability of a label sequence conditioned on the input.
  - ❑ CRFs have previously been used for several natural language processing tasks such as named entity recognition, POS tagging etc.



## Problem Formulation

- ❖ Discussion thread  $T$  is composed of message posts:  
 $T = \langle M_1, M_2, \dots, M_n \rangle$
- ❖ Each message  $M_i$  is composed of sentences:  
 $M_i = \langle m_{i1}, m_{i2}, \dots, m_{is} \rangle$
- ❖ Each sentence  $m_{ij}$  is composed of phrases or n-grams :  
 $m_{ij} = \langle p_{ij1}, p_{ij2}, \dots, p_{ijq} \rangle$
- ❖  $M_i' \subset M_i$  be the set of all problem, resolution related sentences in  $M_i$ .
- ❖  $m_{ij}' \subset m_{ij}$  be the set of all problem, resolution related phrases in  $m_{ij}$ .
- ❖ Tasks: for each  $T$ , find:
  - ❑  $m_{ij} \in M_i'$  for sentence level classification
  - ❑  $p_{ijk} \in m_{ij}'$  for phrase level classification
  - ❑ Classify each as problem or resolution related





## Sequence Labeling using CRFs

- ❖  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$  : input sequence
- ❖  $\mathbf{y} = \{y_1, y_2, \dots, y_m\}$  : label sequence
- ❖ CRF models the probability of label sequence as:

$$P(\mathbf{y} | \mathbf{x}, \mathbf{w}) = \frac{\exp(\sum_j w_j f_j(\mathbf{y}, \mathbf{x}))}{Z(\mathbf{x})}$$

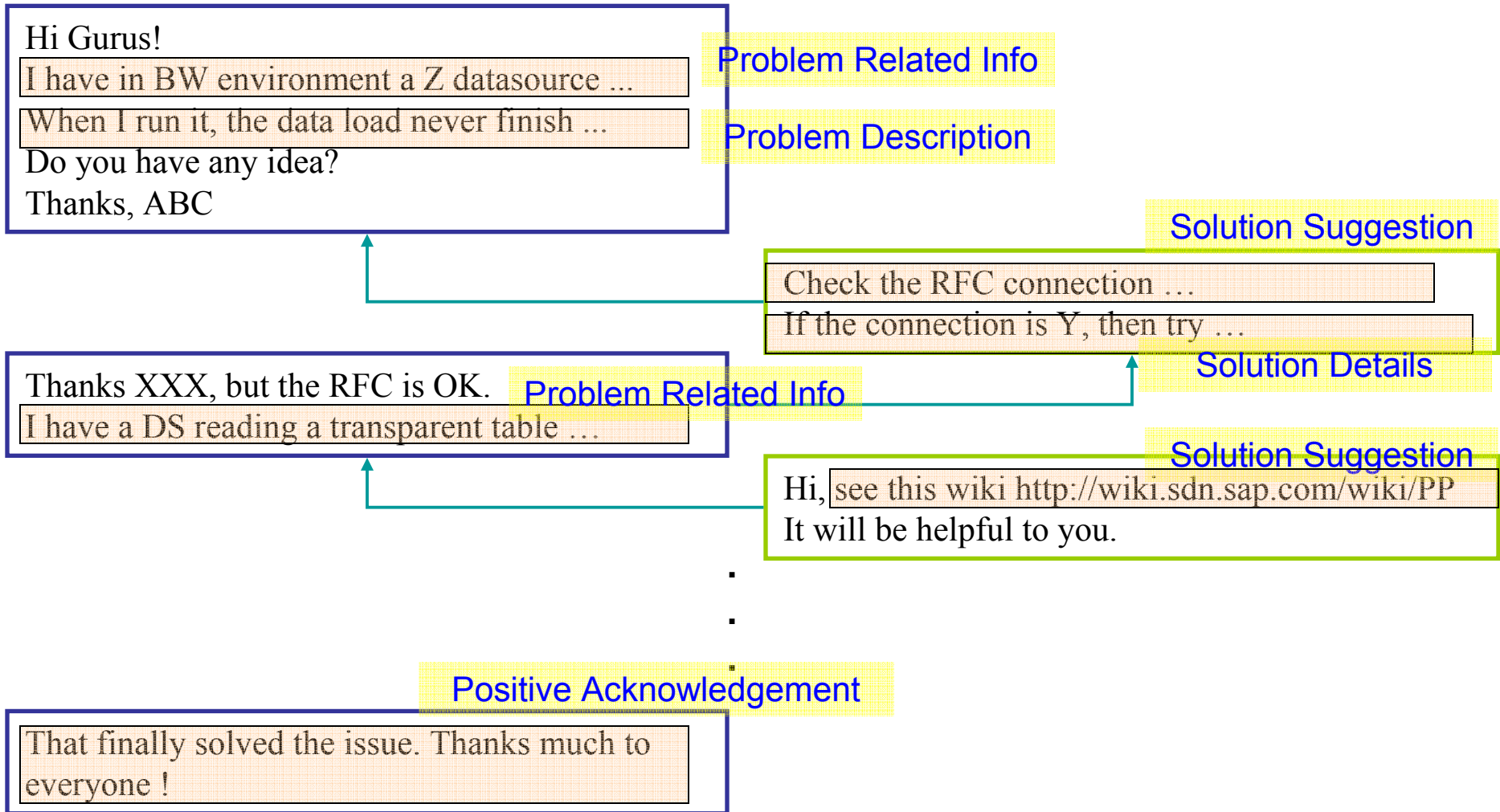
where,

- $f_j$  is the  $j^{\text{th}}$  feature function
- $w_j$  is the weight for  $f_j$  (learned during training)
- $Z(\mathbf{x})$  is a normalization factor

$$\mathbf{y}^* = \operatorname{argmax}_{\mathbf{y}} P(\mathbf{y} | \mathbf{x}, \mathbf{w})$$



## Example Thread





## Labels

- ❖ **PROBLEM:** The problem being addressed in the thread
  - ❑ e.g. As soon as we added SSL encryption it did not work anymore.
- ❖ **RESOLUTION STEP:** Steps leading to problem solution.
  - ❑ e.g. Connect to port 504 for instance number 00
- ❖ **IGNORE:** Remaining part of text which is not relevant to the problem or resolution.
  - ❑ e.g. Have a best day ahead.



## Features

### ❖ Words

- ❑ Captures domain information
- ❑ Baseline for sentence and phrase labeling

### ❖ POS Tags

- ❑ Part-Of-Speech tags add information about grammatical structure and word category
- ❑ Included: Noun, Verb, Adverb and Adjective forms
- ❑ Used Stanford Log-Linear POS Tagger



## Features – contd.

### ❖ **Normalized position of posts**

- ❑ Index of the post normalized wrt the total length of the thread
- ❑ Encoded as ***begin***, ***middle*** and ***end***.
- ❑ Observed trends of:
  - problem description occurring at the beginning of the thread.
  - resolution information typically towards the middle or end.



## Experimental Setup

- ❖ SAP Community Network Forum (<http://forums.sdn.sap.com>)
  - ❑ Discussion site for developers and customers of SAP products
- ❖ Manual annotation using Stanford Manual Annotation Tool
- ❖ CRF implementation – MALLET toolkit



## Experimental Setup - Data Set

### ❖ Sentence Level:

	<b>Train</b>	<b>Test</b>
<b># threads</b>	34	45
<b># sentences</b>	977	1057
<b>Problem</b>	221	174
<b>Resolution Step</b>	268	139

### ❖ Phrase Level:

	<b>Train</b>	<b>Test</b>
<b># phrases</b>	395	168



# Results

❖ Sentence vs. Phrase level : Baselines

Chunking Level	Problem			Resolution Step		
	Precision	Recall	F1	Precision	Recall	F1
<b>Sentence</b>	57.2	59.2	58.2	41.5	61.2	49.5
<b>Phrase</b>	40.0	15.4	22.2	57.1	27.9	37.5

- Phrase level classifier suffers from very low recall values.





## Results – contd.

- ❖ Sentence Level : Accuracy values for different features and combinations

Features	Problem			Resolution Step		
	Precision	Recall	F1	Precision	Recall	F1
Baseline	57.2	59.2	58.2	41.5	61.2	49.5
Baseline + Normalization	61.6	67.8	64.6	40.5	76.3	52.9
Baseline + POS tags	61.5	70.7	65.8	49.7	69.1	57.8
Baseline + Normalization + POS tags	61.0	71.8	66.0	51.6	68.3	58.8



## Conclusions

- ❖ Discussion forums provide an excellent source of information.
- ❖ Studied the basic structure and information flow pattern of a thread.
- ❖ Modeled the task of extracting problem and resolution information as a sequence labeling task.
- ❖ Used CRFs for sentence and phrase levels labeling
- ❖ Studied the effect of using different characteristics of the discussion, on the extraction process



## Conclusions – contd.

- ❖ Sentence level classifier achieves much better performance than phrase level:
  - ❑ 162% improvement for problem extraction
  - ❑ 32% improvement for resolution information extraction
- ❖ Features induced for the sentence level classifier improves the performance, compared to baseline:
  - ❑ 13% improvement for problem extraction
  - ❑ 19% improvement for resolution information extraction



## Future Work

- ❖ Induce more features and study their effect on the extraction process:
  - ❑ Meta-data associated with the threads:
    - Author and the corresponding level of authority
  - ❑ Detect frequent n-grams (phrases) associated with questions and answers
  - ❑ Utilize domain ontology
  
- ❖ Study the usefulness of the extracted data in real world scenarios:
  - ❑ Help desk
  - ❑ Question answering system



## References

- ❖ J. Lafferty, A. McCallum, and F. Pereira. **Conditional random fields: Probabilistic models for segmenting and labeling sequence data**. In *Proc. of ICML, International Conference on Machine Learning*, 2001.
  - ❑ CRFs are probabilistic discriminative models for segmenting and labeling sequence data.
  - ❑ They predict the probability of a label sequence conditioned on the input
  - ❑ Can capture long-range dependencies in the data.
  - ❑ CRFs have previously been used for several natural language processing tasks such as named entity recognition, POS tagging etc.
- ❖ Andrew McCallum. **Mallet: A machine learning for language toolkit**. <http://mallet.cs.umass.edu>, 2002.
- ❖ Stanford Manual Annotation Tool. <http://nlp.stanford.edu/software/stanford-manual-annotation-tool-2004-05-16.tar.gz>



## Related Work

- ❖ V. R. Carvalho and W. W. Cohen. **Learning to extract signature and reply lines from email.** In *Proc. Conference on Email and Anti-Spam*, 2004.
  - Uses CRF and SVM to extract signature and reply lines from emails
- ❖ G. Cong, L. Wang, C. Y. Lin, Y. I. Song, and Y. Sun. **Finding question-answer pairs from online forums.** In *Proc. SIGIR*, 2008.
  - Find question – answer pairs from forum discussions
  - use classifier for questions, unsupervised graph based propagation method for answers
- ❖ S. Ding, G. Cong, C. Y. Lin, and X. Zhu. **Using conditional random fields to extract contexts and answers of questions from online forums.** In *Proc. ACL-HLT, Association for Computational Linguistics*, 2008.
  - Assumes that questions have been identified
  - Applies linear CRFs to extract answers and their contexts from TripAdvisor threads.