软件缺陷报告摘要

报告人：李晓晨
大连理工大学

目录

软件缺陷报告摘要

- 背景：缺陷报告摘要的必要性
- 动机：特征构造在缺陷报告摘要中的地位
- 基于众包的摘要：利用众包的方式构造有效特征
- 基于深度学习的摘要：利用深度学习自动挖掘摘要据的权重

背景

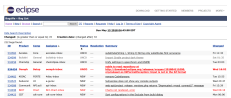
软件工件摘要

> 8,000万 项目

> 300万 应用

> 500万 帖子

持续增长的软件工件促使大量工作对**软件工件摘要**展开研究



源代码摘要

- Haiduc et al. 2010
- Moreno et al. 2013
- Sridhara et al. 2010
- McBurney and McMillan 2011

缺陷报告摘要

- Rastkarude et al. 2014
- Czarnecki et al. 2012
- Mani et al. 2012
- Jiang et al. 2017

源代码注释摘要

- Rastkar et al. 2011
- Ying and Robillard 2013

开发者活动摘要

- Treude et al. 2015

背景

缺陷报告是持续增长的

eclipse DOWNLOAD GETTING STARTED MEMBERS PROJECTS MORE

Bugzilla – Bug List

Home | New | Browse | Search | Search [?] | Reports | Requests | Help | Log In | Terms of Use | Copyright Agent

Sun May 13 2018 04:43:58 EDT

Hide Search Description

Changed: (is greater than or equal to) 7d **Creation date:** (changed after) 7d

212 bugs found.

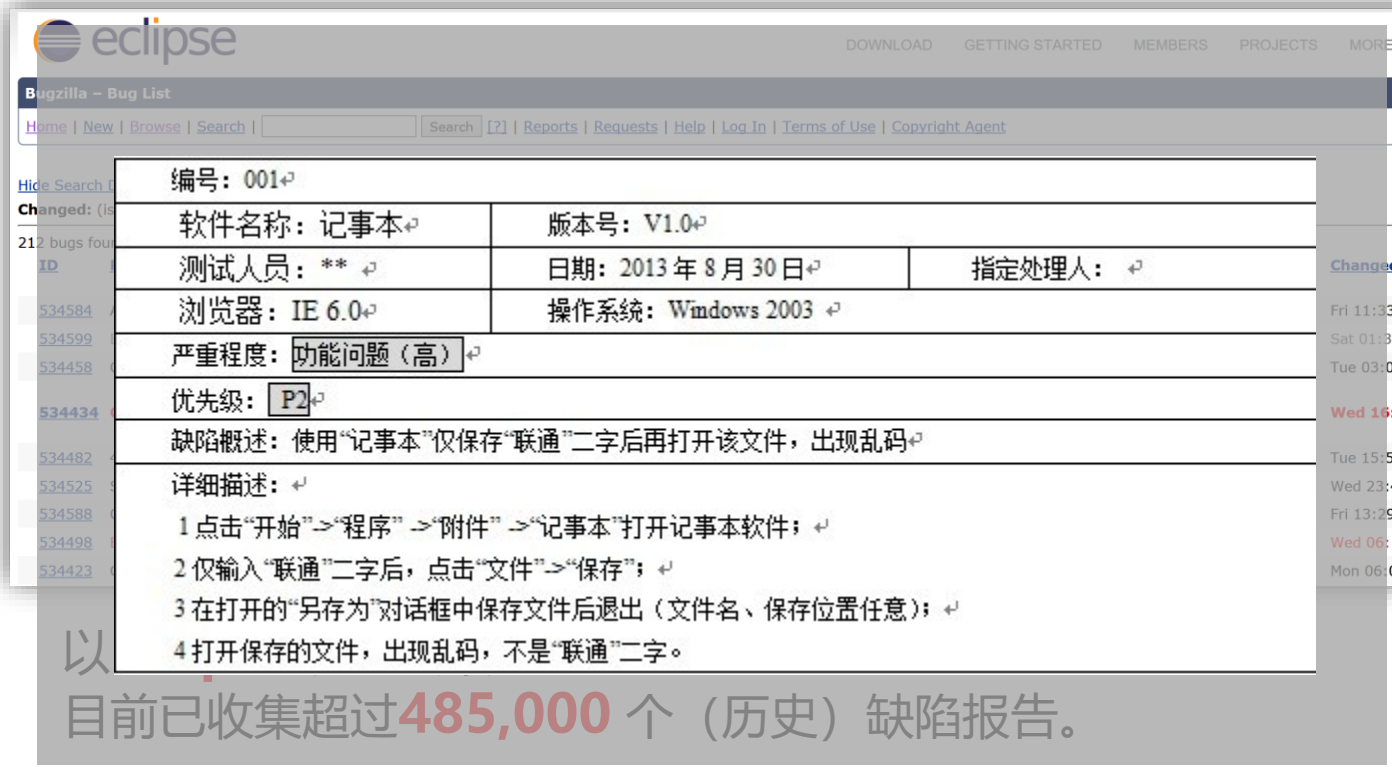
ID	Product	Comp	Assignee ▲	Status ▲	Resolution	Summary	Changed
534584	Acceleo	Core	accceleo-inbox	UNCO	---	substitute(String r, String t):String only substitutes first occurance	Fri 11:33
534599	Buildshi	General	buildship-inbox	UNCO	---	Import Gradle project dark theme	Sat 01:35
534458	Chemclip	General	chemclipse-inbox	UNCO	---	Unify settings in openchrom and KNIME	Tue 03:03
534434	**Oomph**	**Setup**	**oomph-inbox**	**UNCO**	**---**	**nable to read repository at http://download.eclipse.org/releases/oxygen/201804111000. org.tukaani.xz.XZFormatException: Input is not in the XZ format**	**Wed 16:02**
534482	4DIAC	FORTE	4diac-inbox	NEW	---	remove Codntionsync	Tue 15:55
534525	Subversi	UI	a.gurov	NEW	---	Subversise does not show any console outputs	Wed 23:47
534588	Communit	API.ecli	api-inbox	NEW	---	web-api/project_release_versions.php returns "Deprecated: mysql_connect()" message	Fri 13:29
534498	BIRT	Chart	Birt-Chart-inbox	NEW	---	Charts not rendered, error	Wed 06:18
534423	CDT	cdt-core	cdt-core-inbox	NEW	---	Sort configurations in the Exclude from build dialog	Mon 06:06

以**Eclipse**缺陷仓库为例

目前已收集超过**485,000** 个（历史）缺陷报告。

背景

缺陷报告是持续增长的



论软件测试实施. https://www.tuicool.com/articles/a2aYNn

背景

软件项目相关人员会频繁参考历史bug报告

开发者

在修复新bug时会参考相似的历史缺陷报告来**寻找可能的解决方案**

在提交bug报告之前会要求阅读历史相似bug报告以**避免提交重复bug报告**

软件用户和测试人员

背景

软件项目相关人员会频繁参考历史bug报告

开发者

软件用户和测试人员

在修复新bug时会参考相似的历史缺陷报告来**寻找可能的解**

在提交bug报告之前会要求阅读历史相似bug报告以**避免提交**

如果用户仅仅参考**10个历史bug报告**，
他/她平均需要阅读**600个句子**

背景

Bug报告摘要

Bug 170801 - Converting image from grayscale to black&white is painfully slow

Status: RESOLVED FIXED
Product: GIMP
Component: General
Version: 2.2.x
Hardware: Other All

Reported: 2005-03-18 14:46 UTC by Xuan Baldauf
Modified: 2008-01-15 12:50 UTC (History)
CC List: 1 users (show)
See Also:

Importance: Normal normal
Assigned To: GIMP Bugs

Xuan Baldauf 2005-03-18 14:46:31 UTC Description

1. Open a large grayscale image of your choice (e.g. ….
2. Use "Tools/Color Tools/Threshold" to apply some threshold choosen.
3. Now you have a 8bit grayscale image, which acturally consists only of color values "0" and color values "255". ….

……

This slow speed is not acceptable for interactive image processing, and this slowi… not nessary at all.

……

Manish Singh 2005-03-19 17:48:16 UTC Comment 8

Revision 1.156. ….
if (palette_type == GIMP_WEB_PALETTE ||
palette_type == GIMP_MONO_PALETTE ||

……

Adam D. Moss 2005-03-20 12:26:09 UTC Comment 10

The 'mono' palette option doesn't even bother to start this pre-pass because it could only possibly pay off the extra effort if the entire image is pure black and pure white, which is expected to be a comparatively rare occurance.

Xuân Baldauf 2005-03-20 13:06:07 UTC Comment 11

<quote> The 'mono' palette option doesn't ……
I don't think that this operation is so rare, … and then a "convert to 1bit" operation to actually adjust the internal memory requirements.

Adam D. Moss 2005-03-20 14:01:48 UTC Comment 12

> and then a "convert to 1bit" operation to actually adjust
> the internal memory requirements.
If you mean GIMP's internal memory requirements ….

……

从bug报告的描述和评论中**抽取和高亮**具有较高信息量的句子

Xuan Baldauf 2005-03-18 14:46:31 UTC Description

1. Open a large grayscale image of your choice (e.g. ….
2. Use "Tools/Color Tools/Threshold" to apply some threshold choosen.
3. Now you have a 8bit grayscale image, which acturally consists only of color values "0" and color values "255". ….

……

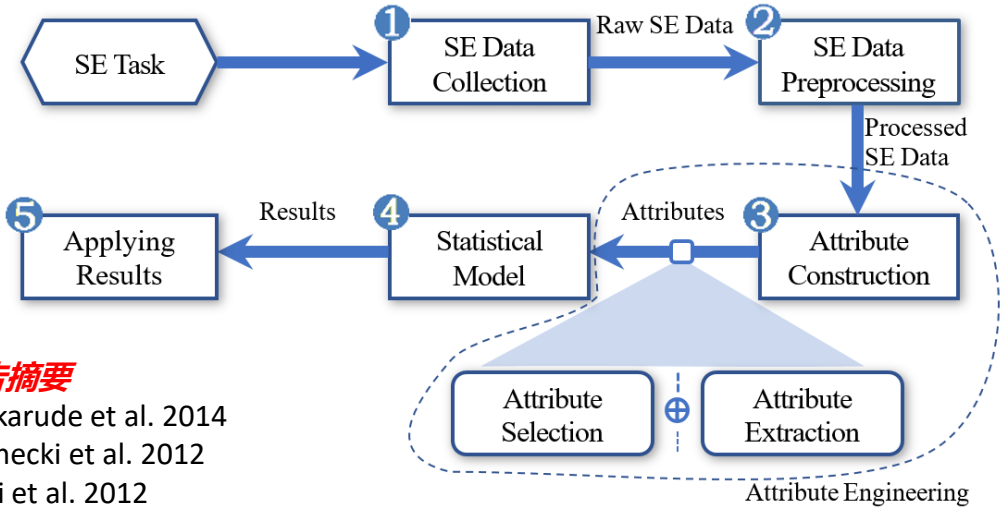
目录

软件缺陷报告摘要

背景	缺陷报告摘要的必要性
动机	特征构造在缺陷报告摘要中的地位
基于众包的摘要	利用众包的方式构造有效特征
基于深度学习的摘要	利用深度学习自动挖掘摘要据的权重

动机

机器学习方法通用过程



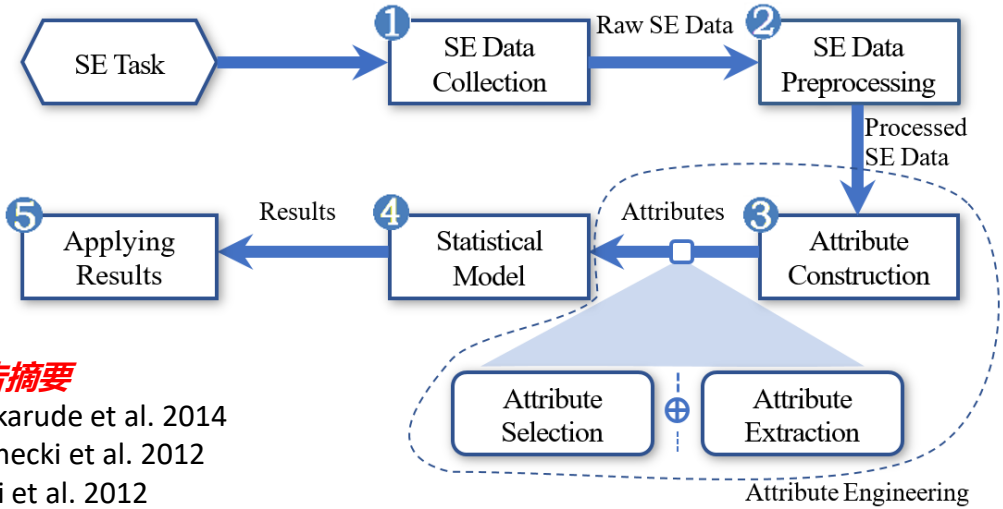
缺陷报告摘要

- Rastkarude et al. 2014
- Czarnecki et al. 2012
- Mani et al. 2012
- Jiang et al. 2017

对于这类方法，**特征至关重要**

动机

机器学习方法通用过程



缺陷报告摘要

- Rastkarude et al. 2014
- Czarnecki et al. 2012
- Mani et al. 2012
- Jiang et al. 2017

对于这类方法，**特征至关重要**

动机

已有研究构造了大量的特征

#	Algorithm	Attribute	Short description
1	BRC	MXS	max *Sprob* score
2		MNS	Mean *Sprob* score
3		SMS	Sum of *Sprob* score
4		MXT	Max *Tprob* score
5		MNT	Mean *Tprob* score
6		SMT	Sum of *Tprob* score
7		TLOC	Position in turn
8		CLOC	Position in conv.
9		SLEN	Word count, globally normalized
10		SLEN2	Word count, locally normalized
11		TPOS1	Time from beg. of conv. to turn
12		TPOS2	Time from turn to end of conv.
13		PPAU	Time btwn. current and prior turn
14		SPAU	Time btwn. current and next turn
15		COS1	Cos. of conv. splits, w/ *Sprob*
16		COS2	Cos. of conv. splits, w/ *Tprob*
17		CENT1	Cos. of sentence & conv., w/ *Sprob*
18		CENT2	Cos. of sentence & conv., w/ *Tprob*
19		PENT	Entropy of conv. up to sentence
20		SENT	Entropy of conv. after the sentence
21		THISENT	Entropy of current sentence
22		DOM	Participator dominance in words
23		BEGAUTH	Is first participator (0/1)
24		CWS	Rough ClueWordScore
25	Centroid MMR	WORD	TF-IDF value of each word
26	Grasshopper DivRank	SENTENCE	Simi. btwn. a sentence and every sentence in the bug report
27	Hurried	TITLE	Simi. btwn. the sentence and title
28		DES	Is the sentence in the description
29		SENTIMENT	The sentiment of a sentence

利用特征可以从多个维度刻画一个句子，进而判断它是否属于摘要据

句子创建的时间
句子单词数量
句子的创建者
句子的情感特性
句子的位置
句子中单词的特殊性
… …

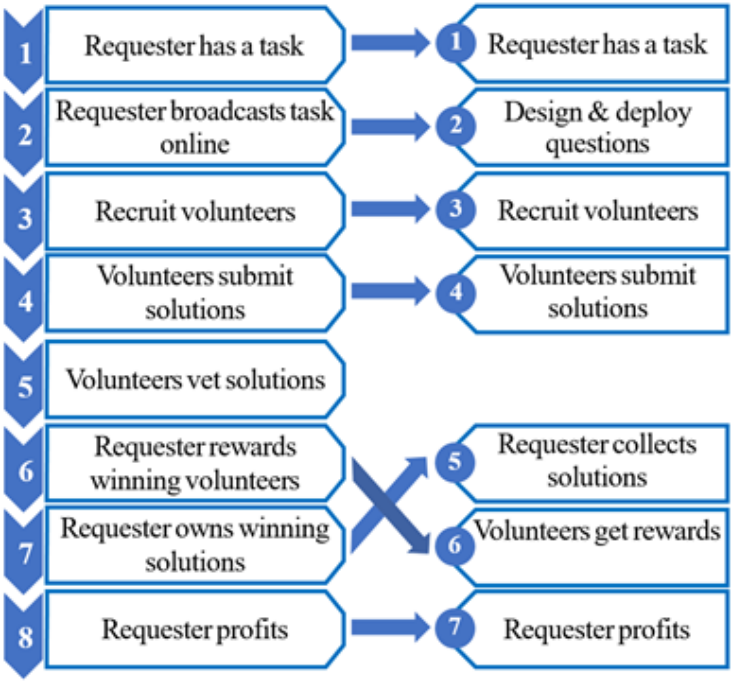
目录

软件缺陷报告摘要

背景	缺陷报告摘要的必要性
动机	特征构造在缺陷报告摘要中的地位
基于众包的摘要	利用众包的方式构造有效特征
基于深度学习的摘要	利用深度学习自动挖掘摘要据的权重

众包特征构建（CA）框架

标准众包过程 → **CA过程**



要求志愿者**人工完成软件工程任务**并**记录进行决策的原因**。

研究者通过**分析这些原因**来为软件工程任务构建特征。

He Jiang, Xiaochen Li, Zhilei Ren, Jifeng Xuan, Zhi Jin. Towards Better Summarizing Bug Reports with Crowdsourcing Elicited Attributes. IEEE Transactions on Reliability, Accepted

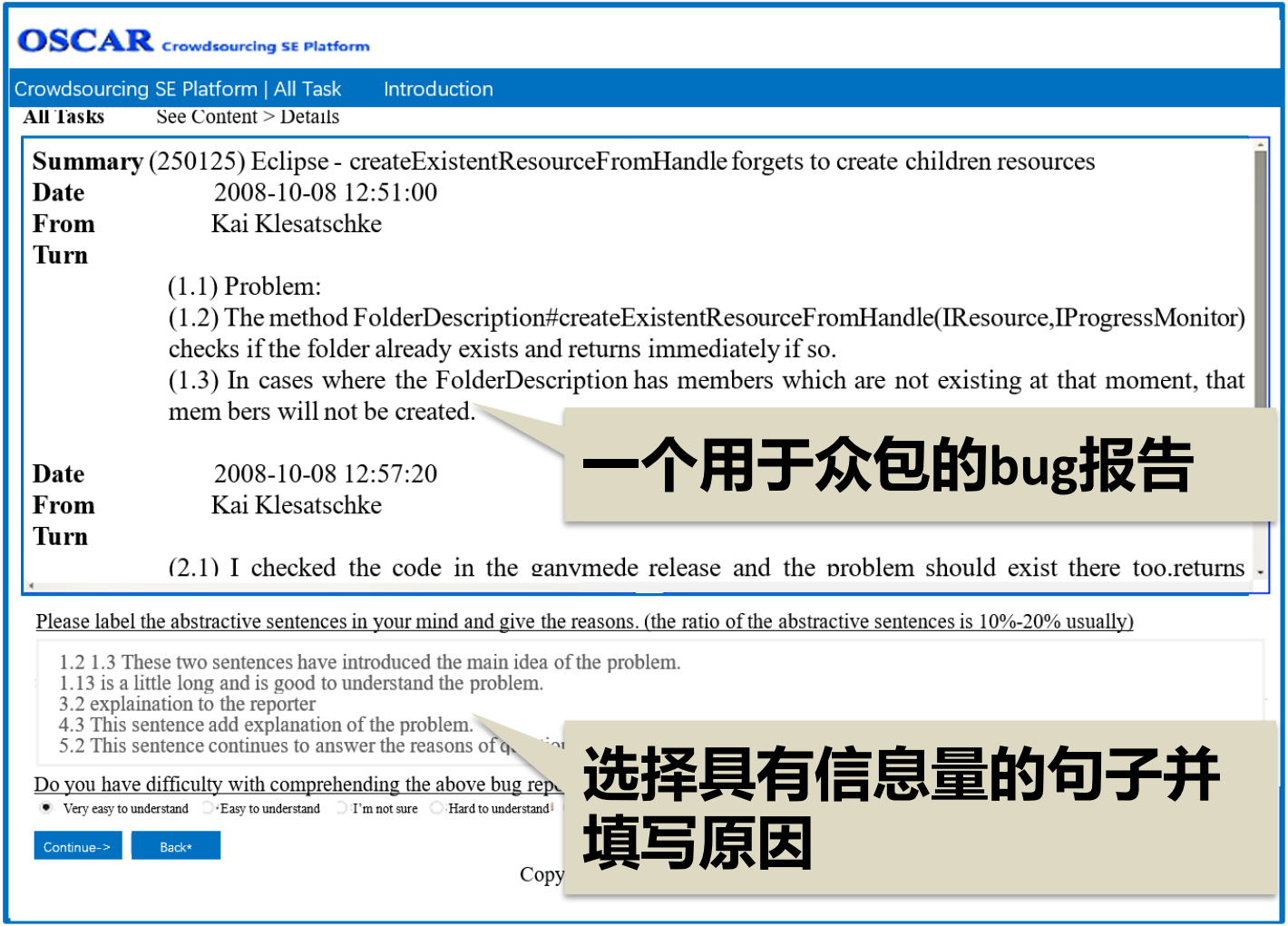
框架

1. 要求志愿者人工完成软件工程任务

1. Requester has a task
2. Design & deploy questions
3. Recruit volunteers
4. Volunteers submit solutions
5. Requester collects solutions
6. Volunteers get rewards
7. Requester profits

1 利用CA方法为**bug报告摘要**任务构建特征

2 从Mozilla, Eclipse, KDE, and Gnome仓库中收集**30个bug报告**用于让志愿者**人工完成摘要任务**.


OSCAR Crowdsourcing SE Platform
Crowdsourcing SE Platform | All Task
Introduction
All Tasks
See Content > Details
Summary (250125) Eclipse - createExistentResourceFromHandle forgets to create children resources
Date 2008-10-08 12:51:00
From Kai Klesatschke
Turn
(1.1) Problem:
(1.2) The method FolderDescription#createExistentResourceFromHandle(IResource,IProgressMonitor) checks if the folder already exists and returns immediately if so.
(1.3) In cases where the FolderDescription has members which are not existing at that moment, that mem bers will not be created.
Date 2008-10-08 12:57:20
From Kai Klesatschke
Turn
(2.1) I checked the code in the ganymede release and the problem should exist there too.returns
一个用于众包的bug报告
Please label the abstractive sentences in your mind and give the reasons. (the ratio of the abstractive sentences is 10%-20% usually)
1.2 1.3 These two sentences have introduced the main idea of the problem.
1.13 is a little long and is good to understand the problem.
3.2 explaination to the reporter
4.3 This sentence add explanation of the problem.
5.2 This sentence continues to answer the reasons of q
选择具有信息量的句子并填写原因
Do you have difficulty with comprehending the above bug rep
Very easy to understand
Easy to understand
I'm not sure
Hard to understand
Continue->
Back*
Copy

框架

2. 要求志愿者记录进行决策的原因

3 **21名志愿者** (给软件工程专业450名学生发送要求邮件)。

4 **1名志愿者** 对 **3个bug报告** 进行摘要。
1个bug报告 被 **2名志愿者** 分析。

5 **19名志愿者**提交了有效的原因。
6 共收集到 **332个**选择特定句子的**原因。**

框架

3. 利用选择句子的原因构造特征

EXAMPLES OF VOLUNTEERS REASONS AND CONSTRUCTED ATTRIBUTES.

Selected Summary Sentence	Volunteer's Reason	Requester's Comment	New Attribute
S1: The idea is that if one prefers to use a web based e-mail account, they should have the option to configure firefox to log into that web based e-mail account and open a com-pose new message window.	R1: the sentence is **long** and the writer puts forward his idea.	long: the length of a sentence may be an attribute	SLEN
	R2: **a rich content**	rich content: we can measure rich content with the sentence length	
S2: Being able to see and sort the history.	R3: **reporter** provides some suggestions	reporter: we can regard the reporter as an attribute	REP
S3: Please add the Mozilla history window in Firefox!	R4: It is **similar with** the sentences in the **bug report**.	similar with bug report: calculate similarity	SWD
	R5: **reporter** provides some suggestions	reporter: be an attribute	REP

志愿者选择具有信息量的摘要句。

志愿者填写选择特定句子的原因。

根据原因构造特征。

框架

启发式构造规则 (HCRs)

志愿者填写的原因
R1: the sentence is ***long*** and the writer puts forward his idea.
R2: ***a rich content***
R3: ***reporter*** provides some suggestions
R4: It is ***similar with*** the sentences in the ***bug report***.
R5: ***reporter*** provides some suggestions

HCR1: 从原因所属句子中抽取候选特征

(a) 高频名词/形容词;

(b) 句子的主语或宾语;

(c) 形容词+名词词组.

框架

启发式构造规则 (HCRs)

"a rich content" 可以理解为 **length**.

分组
R1: *long* (b2)
R2: *a rich content* (b2)
R3, R5: *reporter* (b1)
R4: *similar with … bug report*. (b2)

"similar with, related to, duplicate" 与 **similarity** 相关.

HCR2: 移除无意义候选特征组。

(a) 将候选特征分组

(b) 如果一组候选特征不满足下列要求则删除:

(b1) 候选特征在软件数据的预定义域中出现;

(b2) 候选特征与特定的度量方法相关或可以转换成某种度量。

框架

启发式构造规则 (HCRs)

HCR3: 计算并合并候选特征获得最终特征

*(c1) 候选特征**在相关工作中出现过**: 利用与相关工作相同方法计算特征;*

*(c2) 候选特征是**软件数据预定义的域或者关键词**: 枚举域或关键词可能的状态，将状态值作为特征;*

分组

R3, R5: *reporter*

实例

- “0, 1” 可以表示一个句子是否与特征域相关
- “0, 1, 2” 可以表示bug报告优先级 P1, P2, P3。

框架

启发式构造规则 (HCRs)

HCR3: 计算并合并候选特征获得最终特征

*(c3) 候选特征与 **特定度量相关**: 利用相关的数学定义计算特征;*

*(c4) 剩余特征: 尝试**转换为特定的度量**。*

分组
R1: *long* (length)
R4: *similar with* (similarity)

分组
R2: *a rich content* (length)

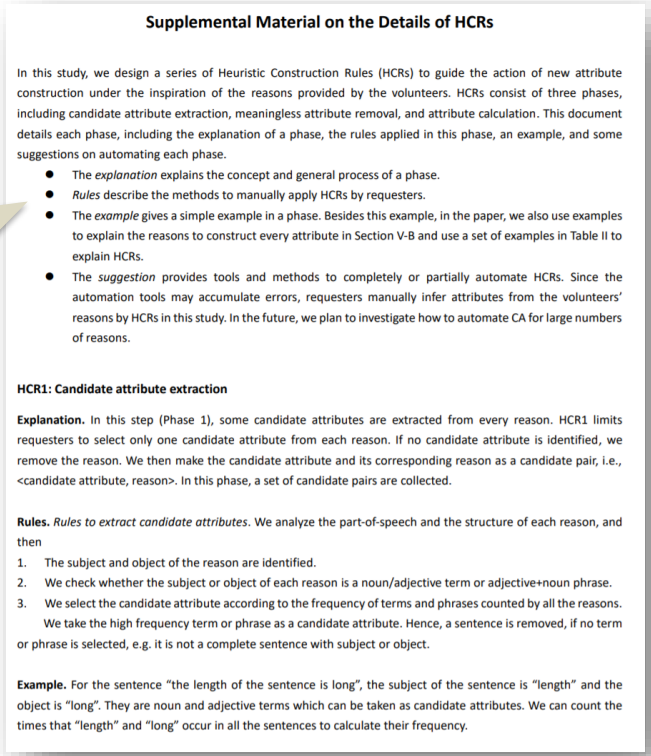
框架

启发式构造规则 (HCRs)

规则详细信息 可以访问网址: http://oscar-lab.org/people/~xcli/open/crowdsourcing/datas/Supplemental_Material_on_the_Details_of_HCRs.pdf

风险

1. 该过程需要研究者背景知识。
2. 它是一个主观的过程。



Supplemental Material on the Details of HCRs

In this study, we design a series of Heuristic Construction Rules (HCRs) to guide the action of new attribute construction under the inspiration of the reasons provided by the volunteers. HCRs consist of three phases, including candidate attribute extraction, meaningless attribute removal, and attribute calculation. This document details each phase, including the explanation of a phase, the rules applied in this phase, an example, and some suggestions on automating each phase.

- The *explanation* explains the concept and general process of a phase.
- *Rules* describe the methods to manually apply HCRs by requesters.
- The *example* gives a simple example in a phase. Besides this example, in the paper, we also use examples to explain the reasons to construct every attribute in Section V-B and use a set of examples in Table II to explain HCRs.
- The *suggestion* provides tools and methods to completely or partially automate HCRs. Since the automation tools may accumulate errors, requesters manually infer attributes from the volunteers' reasons by HCRs in this study. In the future, we plan to investigate how to automate CA for large numbers of reasons.

HCR1: Candidate attribute extraction

Explanation. In this step (Phase 1), some candidate attributes are extracted from every reason. HCR1 limits requesters to select only one candidate attribute from each reason. If no candidate attribute is identified, we remove the reason. We then make the candidate attribute and its corresponding reason as a candidate pair, i.e., <candidate attribute, reason>. In this phase, a set of candidate pairs are collected.

Rules. *Rules to extract candidate attributes*. We analyze the part-of-speech and the structure of each reason, and then

1. The subject and object of the reason are identified.
2. We check whether the subject or object of each reason is a noun/adjective term or adjective+noun phrase.
3. We select the candidate attribute according to the frequency of terms and phrases counted by all the reasons. We take the high frequency term or phrase as a candidate attribute. Hence, a sentence is removed, if no term or phrase is selected, e.g. it is not a complete sentence with subject or object.

Example. For the sentence "the length of the sentence is long", the subject of the sentence is "length" and the object is "long". They are noun and adjective terms which can be taken as candidate attributes. We can count the times that "length" and "long" occur in all the sentences to calculate their frequency.

框架

利用CA构造特征

#	Attribute	Short description
1	SWT	Similarity With the Topic of bug report
2	SWD	Similarity With Description
3	DUP	is the sentence a DUPlicate sentence
4	SLEN	the normalized Sentence LENgth
5	SI	the normalized Sentence Importance
6	SLOC	the normalized Sentence LOCation
7	CLEN	the normalized Comment LENgth
8	DES	is the sentence in the DEScription
9	CCW	does the sentence Contain Certain Words
10	CODE	is the sentence in a piece of CODE snippet
11	REP	is the sentence written by the REPorter

基于CA构造的特征训练 **逻辑回归Logistic Regression** 模型 (**LRCA**)

评价

利用4个指标在SDS数据集上评价

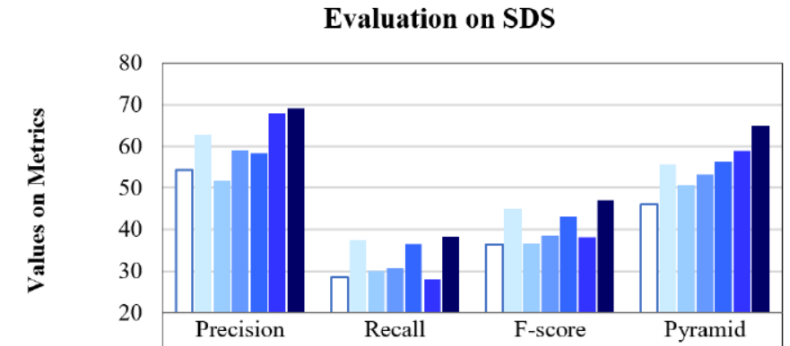
- 评价指标包括 precision, recall, F-score, pyramid

$$
\begin{aligned}
Precision &= Num_{\text{hit}}/Num_{\text{selected}},\\
Recall &= Num_{\text{hit}}/Num_{\text{ExfRef}},\\
F\text{-}score &= \frac{2 * Precision * Recall}{Precision + Recall},\\
Pyramid &= Num_{\text{TotalLinks}}/Num_{\text{MaxLinks}},
\end{aligned}
$$

- SDS数据集包括36个人工标注摘要句的bug报告

评价

LRCA与已有算法对比



	Precision	Recall	F-score	Pyramid
Centroid	54.26	28.61	36.35	46.11
MMR	62.72	37.34	45.06	55.67
Grasshopper	51.83	30.15	36.65	50.63
DivRank	59.03	30.66	38.55	53.19
Hurried	58.41	36.47	43.01	56.32
BRC	67.79	28.16	38.19	58.99
LRCA	69.12	38.27	47.13	64.88

LRCA 超过已有算法。CA构造的特征在进行bug报告摘要时具有优势

Bug报告摘要

- Rastkarude et al. 2014
- Czarnecki et al. 2012
- Mani et al. 2012

目录

软件缺陷报告摘要

背景	缺陷报告摘要的必要性
动机	特征构造在缺陷报告摘要中的地位
基于众包的摘要	利用众包的方式构造有效特征
基于深度学习的摘要	利用深度学习自动挖掘摘要据的权重

分析

缺陷报告特点

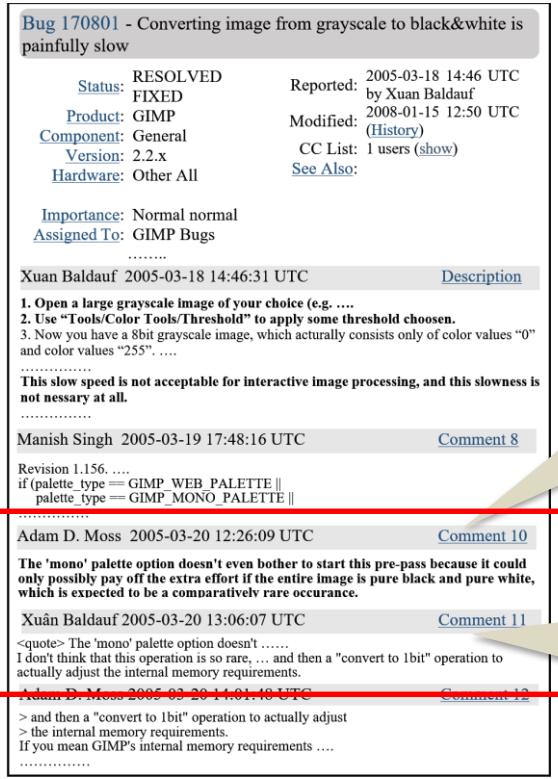
BUG REPORTS are

1. 基于会话的文本带有频繁的**评价行为**
2. 句子包括**丰富的类型**
3. 包含大量的**预定义域 (predefine fields)**

Xiaochen Li, He Jiang, Dong Liu, Zhilei Ren, Ge Li. Unsupervised Deep Bug Report Summarization. In Proceedings of IEEE/ACM ICPC Conference (ICPC'18), Gothenburg, Sweden, May 27-28, 2018.

分析

1. 基于会话的文本带有频繁的评价行为



Bug 170801 - Converting image from grayscale to black&white is painfully slow

Status: RESOLVED FIXED
Product: GIMP
Component: General
Version: 2.2.x
Hardware: Other All

Reported: 2005-03-18 14:46 UTC by Xuan Baldauf
Modified: 2008-01-15 12:50 UTC (History)
CC List: 1 users (show)
See Also:

Importance: Normal normal
Assigned To: GIMP Bugs

……..

Xuan Baldauf 2005-03-18 14:46:31 UTC Description

1. Open a large grayscale image of your choice (e.g. ….
2. Use "Tools/Color Tools/Threshold" to apply some threshold choosen.
3. Now you have a 8bit grayscale image, which acturally consists only of color values "0" and color values "255". ….

……

This slow speed is not acceptable for interactive image processing, and this slowness is not nessary at all.

……

Manish Singh 2005-03-19 17:48:16 UTC Comment 8

Revision 1.156. ….
if (palette_type == GIMP_WEB_PALETTE ||
palette_type == GIMP_MONO_PALETTE ||

……

Adam D. Moss 2005-03-20 12:26:09 UTC Comment 10

The 'mono' palette option doesn't even bother to start this pre-pass because it could only possibly pay off the extra effort if the entire image is pure black and pure white, which is expected to be a comparatively rare occurance.

Xuân Baldauf 2005-03-20 13:06:07 UTC Comment 11

<quote> The 'mono' palette option doesn't ……
I don't think that this operation is so rare, … and then a "convert to 1bit" operation to actually adjust the internal memory requirements.

~~Adam D. Moss 2005-03-20 14:01:48 UTC Comment 12~~

> and then a "convert to 1bit" operation to actually adjust
> the internal memory requirements.
If you mean GIMP's internal memory requirements ….

……

被评价的句子通常会被频繁讨论，**更重要**

Adam D. Moss 2005-03-20 12:26:09 UTC Comment 10

The 'mono' palette option doesn't even bother to star……

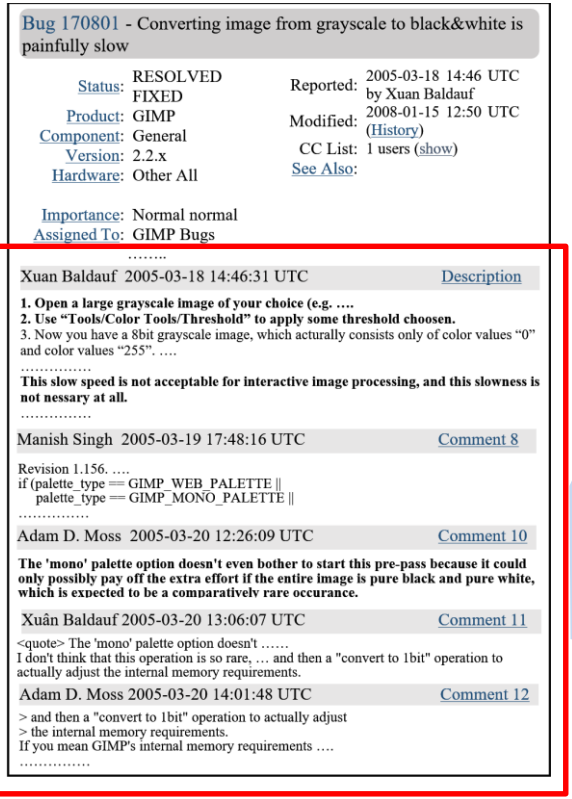
被评价

Xuân Baldauf 2005-03-20 13:06:07 UTC Comment 11

<*quote*> The 'mono' palette option doesn't ……
I don't think that this operation is so rare, …

分析

2. 句子包括丰富的类型



Bug 170801 - Converting image from grayscale to black&white is painfully slow

Status:	RESOLVED FIXED	Reported:	2005-03-18 14:46 UTC by Xuan Baldauf
Product:	GIMP	Modified:	2008-01-15 12:50 UTC (History)
Component:	General	CC List:	1 users (show)
Version:	2.2.x	See Also:	
Hardware:	Other All		
Importance:	Normal normal		
Assigned To:	GIMP Bugs		

……..

Xuan Baldauf 2005-03-18 14:46:31 UTC Description

1. Open a large grayscale image of your choice (e.g. ….
2. Use "Tools/Color Tools/Threshold" to apply some threshold choosen.
3. Now you have a 8bit grayscale image, which acturally consists only of color values "0" and color values "255". ….

……

This slow speed is not acceptable for interactive image processing, and this slowness is not nessary at all.

……

Manish Singh 2005-03-19 17:48:16 UTC Comment 8

Revision 1.156. ….
if (palette_type == GIMP_WEB_PALETTE ||
palette_type == GIMP_MONO_PALETTE ||

……

Adam D. Moss 2005-03-20 12:26:09 UTC Comment 10

The 'mono' palette option doesn't even bother to start this pre-pass because it could only possibly pay off the extra effort if the entire image is pure black and pure white, which is expected to be a comparatively rare occurance.

Xuân Baldauf 2005-03-20 13:06:07 UTC Comment 11

<quote> The 'mono' palette option doesn't ……
I don't think that this operation is so rare, … and then a "convert to 1bit" operation to actually adjust the internal memory requirements.

Adam D. Moss 2005-03-20 14:01:48 UTC Comment 12

> and then a "convert to 1bit" operation to actually adjust
> the internal memory requirements.
If you mean GIMP's internal memory requirements ….

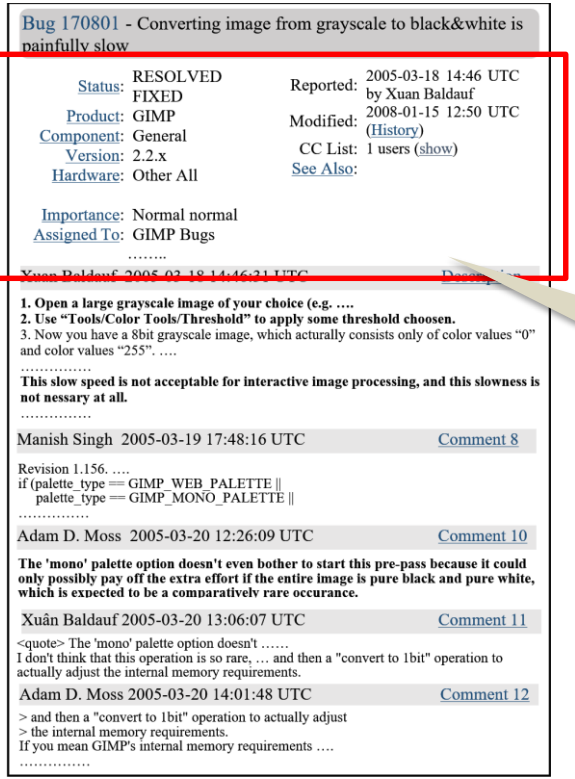
……

reporter's sentences **>**
participators' sentences **>**
software sentences
(`>' means more informative)

- **提交者用自然语言书写的句子**;
- **参与者用自然语言书写的句子**;
- **用软件语言书写的句子** (代码片段和系统消息).

分析

3. 包含大量的预定义域（predefine fields）



Bug 170801 - Converting image from grayscale to black&white is painfully slow

Status:	RESOLVED FIXED	Reported:	2005-03-18 14:46 UTC by Xuan Baldauf
Product:	GIMP	Modified:	2008-01-15 12:50 UTC (History)
Component:	General	CC List:	1 users (show)
Version:	2.2.x	See Also:	
Hardware:	Other All		
Importance:	Normal normal		
Assigned To:	GIMP Bugs		

……..

Xuan Baldauf 2005-03-18 14:46:31 UTC Descrip…

1. Open a large grayscale image of your choice (e.g. ….
2. Use "Tools/Color Tools/Threshold" to apply some threshold choosen.
3. Now you have a 8bit grayscale image, which acturally consists only of color values "0" and color values "255". ….

……

This slow speed is not acceptable for interactive image processing, and this slowness is not nessary at all.

……

Manish Singh 2005-03-19 17:48:16 UTC Comment 8

Revision 1.156. ….
if (palette_type == GIMP_WEB_PALETTE ||
palette_type == GIMP_MONO_PALETTE ||

……

Adam D. Moss 2005-03-20 12:26:09 UTC Comment 10

The 'mono' palette option doesn't even bother to start this pre-pass because it could only possibly pay off the extra effort if the entire image is pure black and pure white, which is expected to be a comparatively rare occurance.

Xuân Baldauf 2005-03-20 13:06:07 UTC Comment 11

<quote> The 'mono' palette option doesn't ……
I don't think that this operation is so rare, … and then a "convert to 1bit" operation to actually adjust the internal memory requirements.

Adam D. Moss 2005-03-20 14:01:48 UTC Comment 12

> and then a "convert to 1bit" operation to actually adjust
> the internal memory requirements.
If you mean GIMP's internal memory requirements ….

……

在预定义域中的句子可能信息量更大

Product: GIMP
Component: General
Version: 2.2.x
Hardware: Other All

分析

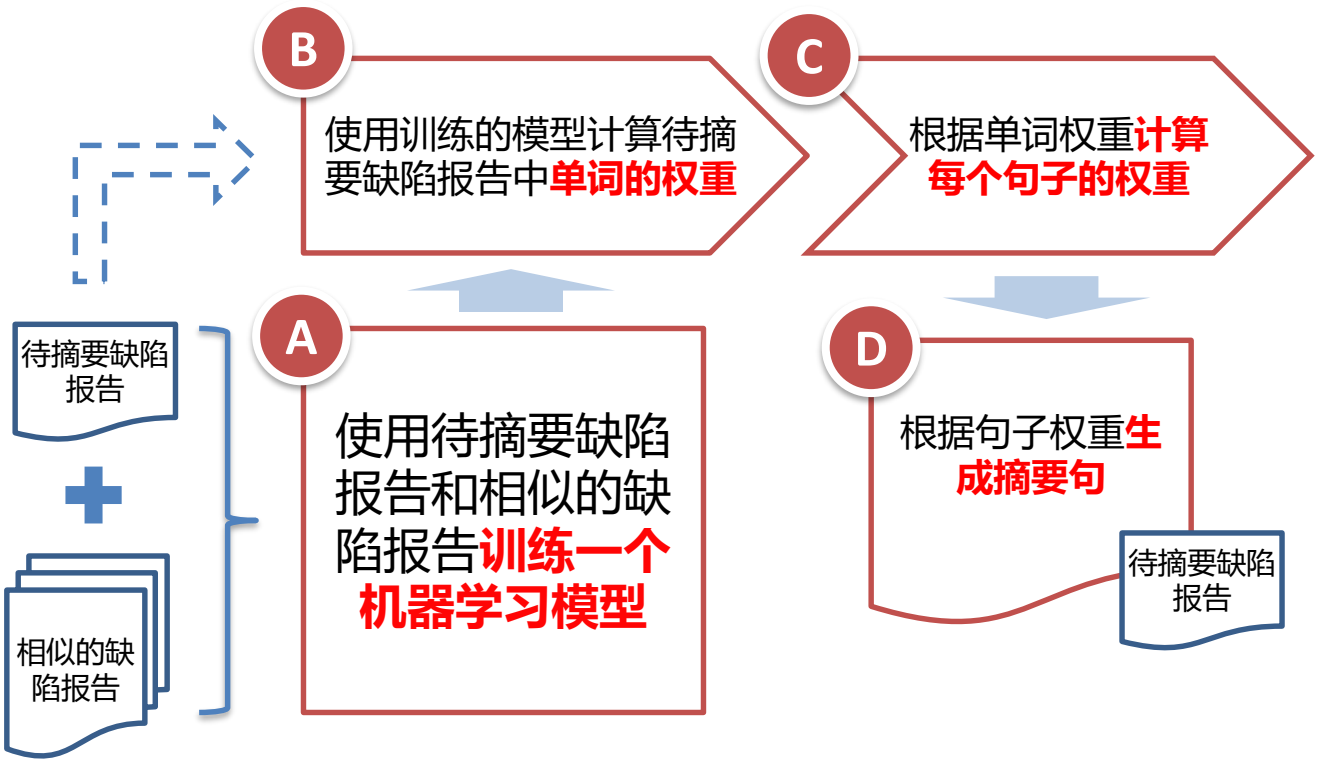
缺陷报告特点

1. 基于会话的文本带有频繁的**评价行为**
2. 句子包括**丰富的类型**
3. 包含大量的**预定义域 (predefine fields)**

在考虑这些特性的情况下进行更准确的摘要

框架

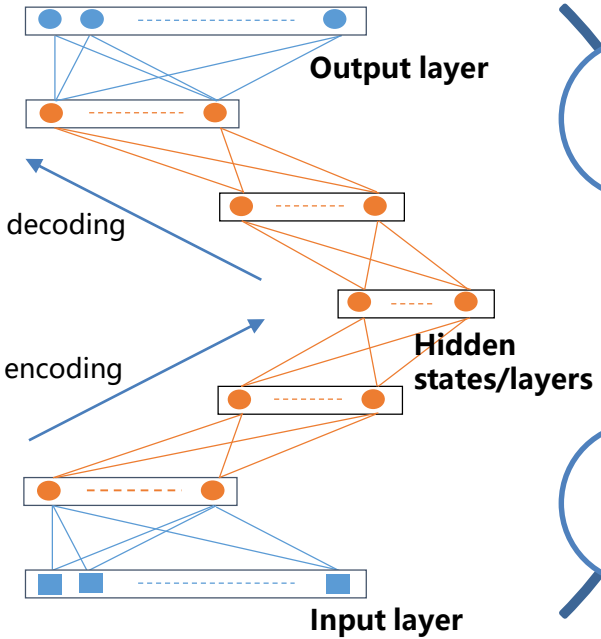
深度学习摘要框架


B
使用训练的模型计算待摘要缺陷报告中单词的权重
C
根据单词权重计算每个句子的权重
待摘要缺陷报告
相似的缺陷报告
A
使用待摘要缺陷报告和相似的缺陷报告训练一个机器学习模型
D
根据句子权重生成摘要句
待摘要缺陷报告

框架

我们选择AutoEncoder（自编码机）作为机器学习模型

一个典型的AutoEncoder框架



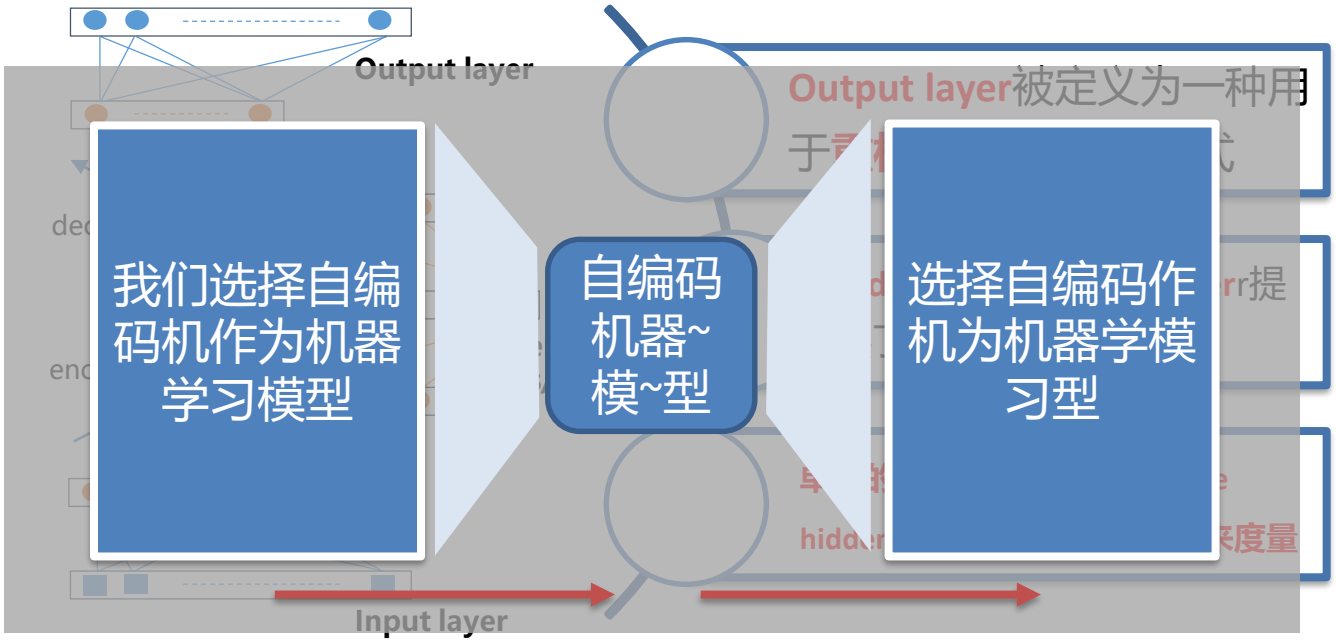
为什么选择AutoEncoder?

- **Output layer**被定义为一种用于**重构input layer**的模式
- **Hidden state**为**input layer**r提供了一种**压缩表示**
- **单词的权重**可以通过它们为**the hidden states贡献了多少信息来度量**

框架

我们选择AutoEncoder（自编码机）作为机器学习模型

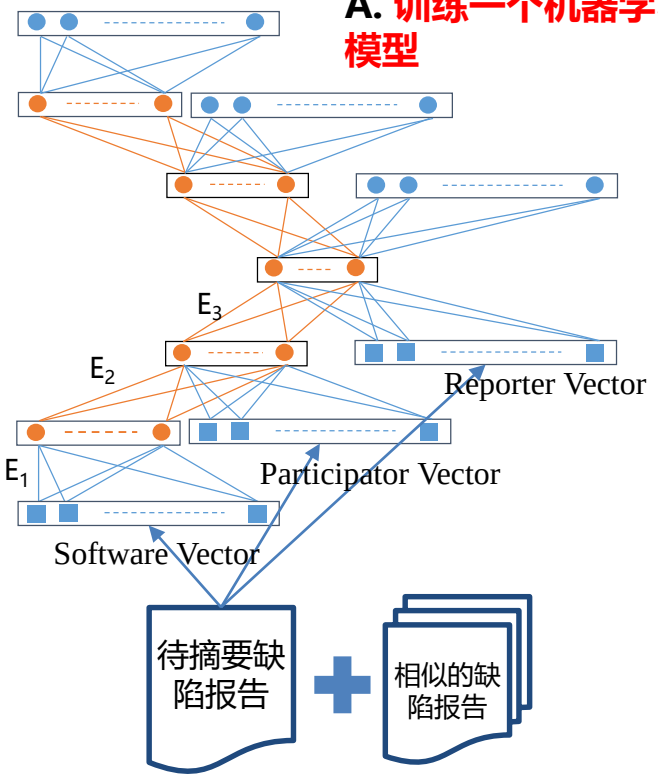
一个典型的AutoEncoder框架



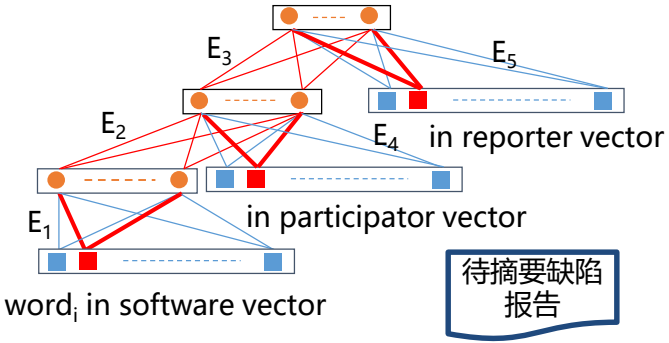
为什么选择AutoEncoder?

35

深度学习框架(DeepSum)



B. 根据训练的参数, e.g., E_1, E_2, E_3, E_4, E_5, **计算单词权重**



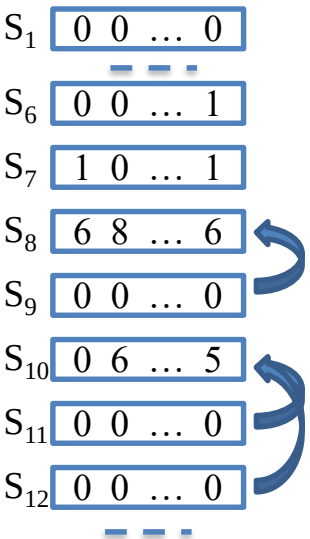
Weight of $word_i$ =

$E_3(E_2(E_1(word_i \text{ in software vector})))$
$+ E_3(E_4(word_i \text{ in participator vector}))$
$+ E_5(word_i \text{ in reporter vector}))$

深度学习框架(DeepSum)

C. 计算每个句子的得分. 句子得分 =

$$\sum word\ weight * word\ frequency$$



D. 生成摘要句, 依据下面两点选择一组句子集合s_{select}作为摘要句:

(1) 最大化s_{select}中句子的得分之和;

(2) S_{select}中句子的粽长度< 长度阈值 (待摘要缺陷报告的25%)

我们利用**动态规划**选择摘要句集合. 摘要的**长度**大概是待摘要缺陷报告的**¼**并且所有句子得分之和**较高**.

评价

在两个数据集上用六个指标进行评价

- 评价指标包括precision, recall, F-score, pyramid, R1, R2

$$Precision = Num_{\text{hit}}/Num_{\text{selected}},$$
$$Recall = Num_{\text{hit}}/Num_{\text{ExfRef}},$$
$$F\text{-}score = \frac{2 * Precision * Recall}{Precision + Recall},$$
$$Pyramid = Num_{\text{TotalLinks}}/Num_{\text{MaxLinks}},$$
$$Rouge\text{-}n = \frac{\sum_{s \in \text{AbsRef}} \sum_{gram_n \in s} Count_{\text{match}}(gram_n)}{\sum_{s \in \text{AbsRef}} \sum_{gram_n \in s} Count(gram_n)},$$

- 利用SDS和ADS两个人工标注的数据集进行评价，共包括36+96个缺陷报告

评价

将DeepSum和其它已有算法进行比较

	Precision	Recall	F-score	Pyramid	R-1	R-2
BRC_{LOO}	0.570	0.350	0.400	0.630	0.521	0.140
BRC_{TFCV}	0.524	0.321	0.362	0.580	0.493	0.130
ACS_{LOO}	0.595	0.337	0.400	0.604	0.516	0.135
ACS_{TFCV}	0.562	0.310	0.359	0.572	0.488	0.126
Centroid	0.536	0.269	0.343	0.460	0.471	0.126
MMR	0.617	0.353	0.429	0.551	0.498	0.145
Grasshopper	0.525	0.300	0.368	0.521	0.505	0.135
DivRank	0.591	0.301	0.378	0.546	0.500	0.139
Hurried	0.710	0.390	0.410	0.710	0.525	0.153
DeepSsum	0.621	**0.388**	**0.462**	0.624	**0.563**	**0.177**

DeepSum在多个指标上的性能超过了其它的算法

	Precision	Recall	F-score	Pyramid	R-1	R-2
BRC_{LOO}	0.568	0.350	0.412	0.659	0.517	0.201
BRC_{TFCV}	0.528	0.314	0.388	0.620	0.492	0.180
ACS_{LOO}	0.605	0.391	0.452	0.671	0.546	0.235
ACS_{TFCV}	0.556	0.343	0.400	0.625	0.520	0.211
Centroid	0.488	0.280	0.337	0.561	0.473	0.183
MMR	0.505	0.356	0.395	0.585	0.503	0.206
Grasshopper	0.446	0.337	0.362	0.548	0.504	0.201
DivRank	0.445	0.282	0.325	0.545	0.498	0.202
Hurried	0.580	0.349	0.418	0.637	0.544	0.241
DeepSsum	**0.606**	**0.394**	**0.457**	**0.681**	**0.553**	**0.249**

缺陷报告摘要

- Rastkarude et al. 2014
- Czarnecki et al. 2012
- Mani et al. 2012
- Jiang et al. 2017

结论

- 介绍Crowd-Attribute (CA)方法。它利用众包群体生成的数据来为bug报告摘要任务**系统的推测特征。**
  - CA还被应用于下列研究中
  - Nazar et al. Source code fragment summarization with small-scale crowdsourcing based features. FCS, 2016, 10(3): 504-517.
  - Jiang et al. What Makes a Good App Description? Internetware '2014, pp.45-53.

- 介绍了一种**无监督的深度学习方法**自动寻找缺陷报告中的摘要句。

谢　谢

软件缺陷报告摘要

报告人：李晓晨
大连理工大学