Large-Scale Patch Recommendation at Alibaba

Xindong Zhang[1], Chenguang Zhu[2], Yi Li[3], Jianmei Guo[1], Lihua Liu[1], and Haobo Gu[1]

1. Alibaba Group 2. University of Texas at Austin 3. Nanyang Technological University

Motivation

50% time

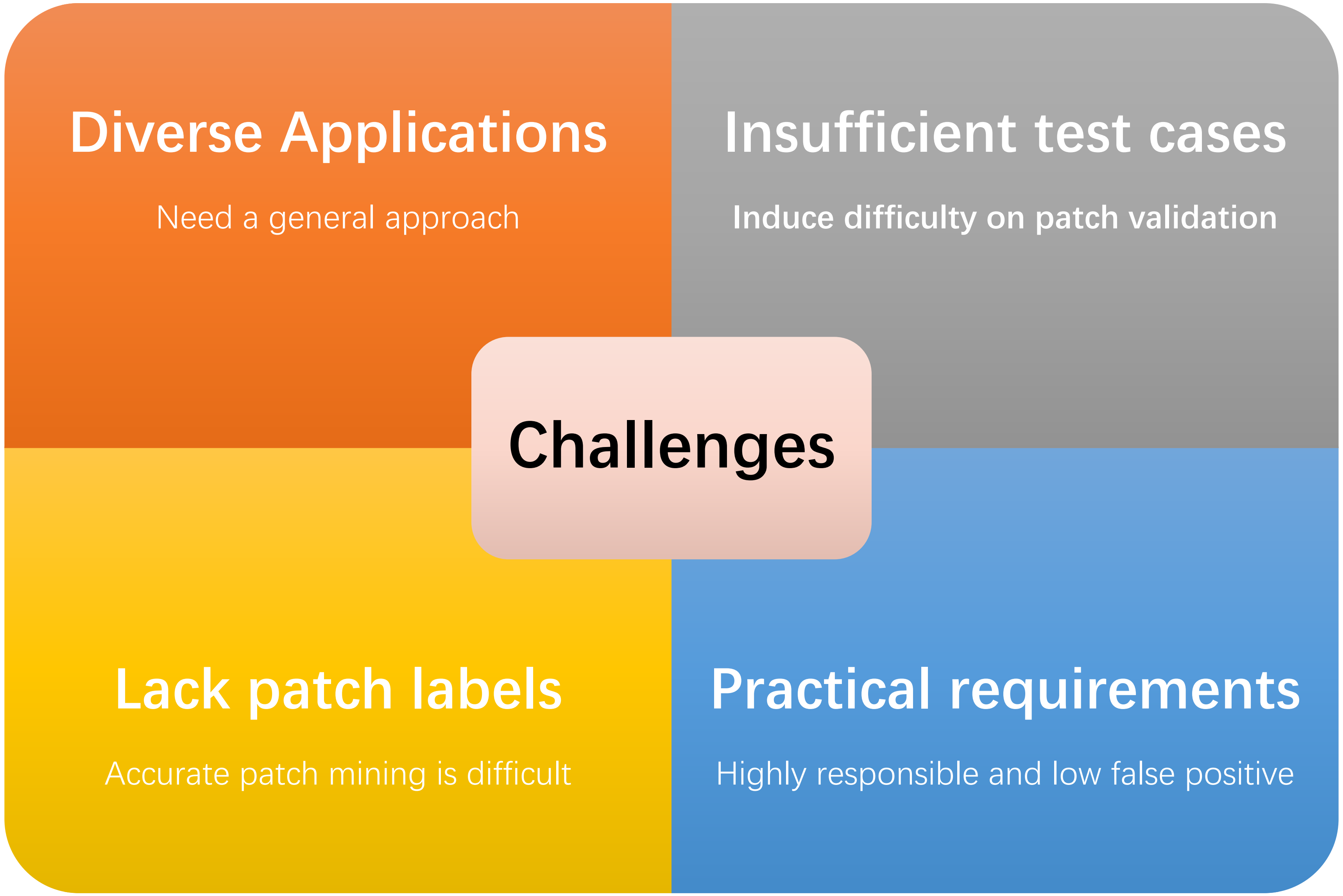
On average, 49.9% of software developers' time has been spent in debugging

50% cost

About half of the development costs are associated with debugging and patching

Automated patch recommendation can significantly reduce developers' debugging efforts and the overall development costs

Challenges



Diverse Applications

Need a general approach

Insufficient test cases

Induce difficulty on patch validation

Challenges

Lack patch labels

Accurate patch mining is difficult

Practical requirements

Highly responsible and low false positive

Solution

Diverse Applications

Patches are mined from internal codebase using generic features

Insufficient test cases

Independent of test cases and use developers' feedback to validate and improve

PRECFIX

Lack patch labels

Automatically mines bug and fix templates from historical changes

Practical requirements

Guarantee high responsiveness (scale of ms) and low false positive (22% and lower)

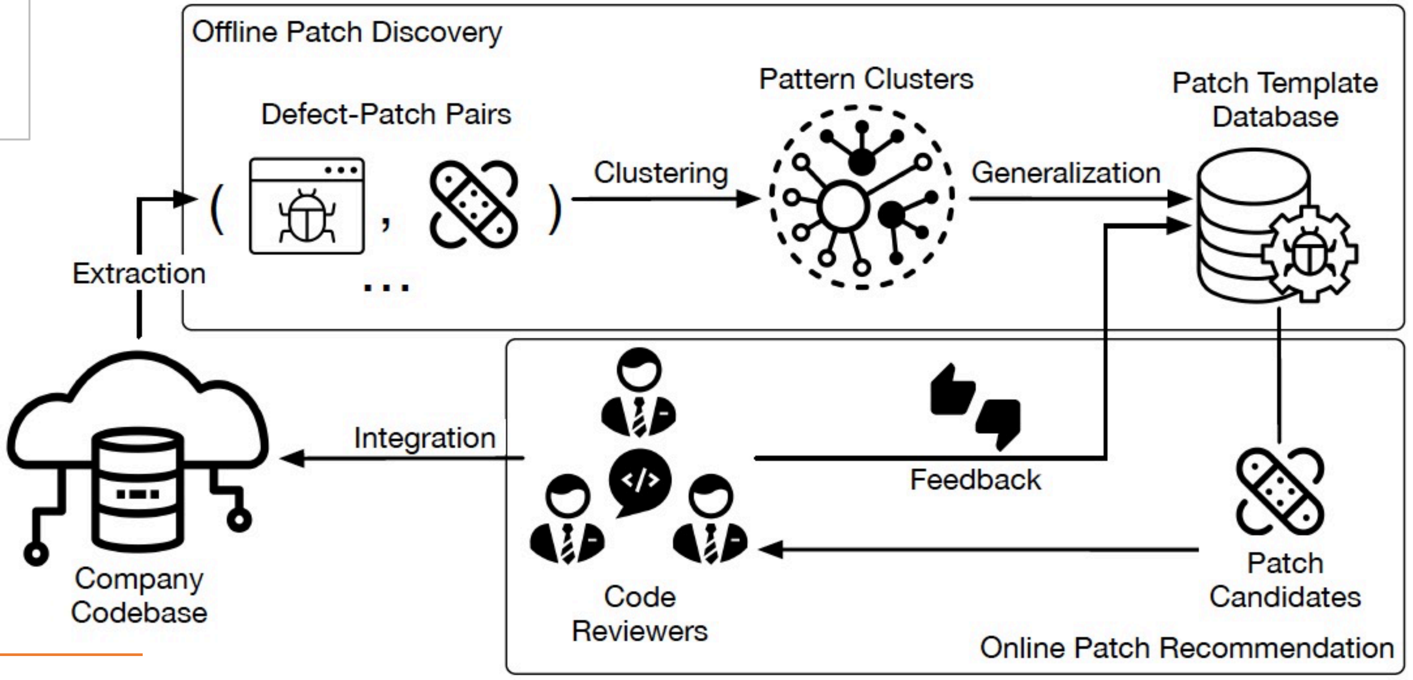
PRECFIX

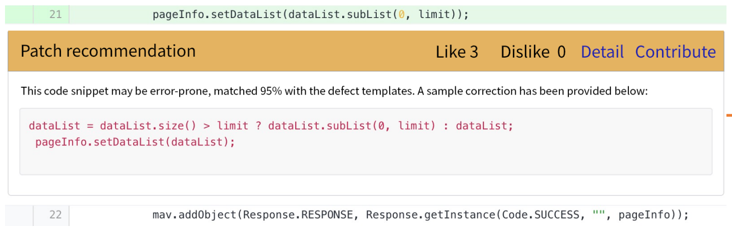
阿里云

Commit
fix#723 NPE check
author: Jack
++++
- - - - -
++++

- Commit message contains fix intentions
- 75% bug-fixing commits have such pattern:
  Delete bug snippet & Add patch snippet

- Clustering Algorithm：DBSCAN
- Clustering Strategy： Both defect & patch snippets
- Optimization：Simhash-KDTree, API sequence
- Similarity Comparison：Levenstein + Jaccard





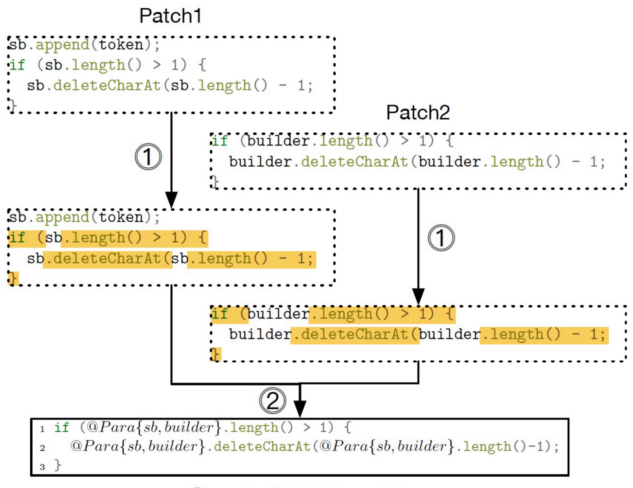
```
pageInfo.setDataList(dataList.subList(0, limit));
```

Patch recommendation   Like 3   Dislike 0   Detail   Contribute

This code snippet may be error-prone, matched 95% with the defect templates. A sample correction has been provided below:

```
dataList = dataList.size() > limit ? dataList.subList(0, limit) : dataList;
 pageInfo.setDataList(dataList);
```

```
mav.addObject(Response.RESPONSE, Response.getInstance(Code.SUCCESS, "", pageInfo));
```



```
sb.append(token);
if (sb.length() > 1) {
  sb.deleteCharAt(sb.length() - 1;
}
```

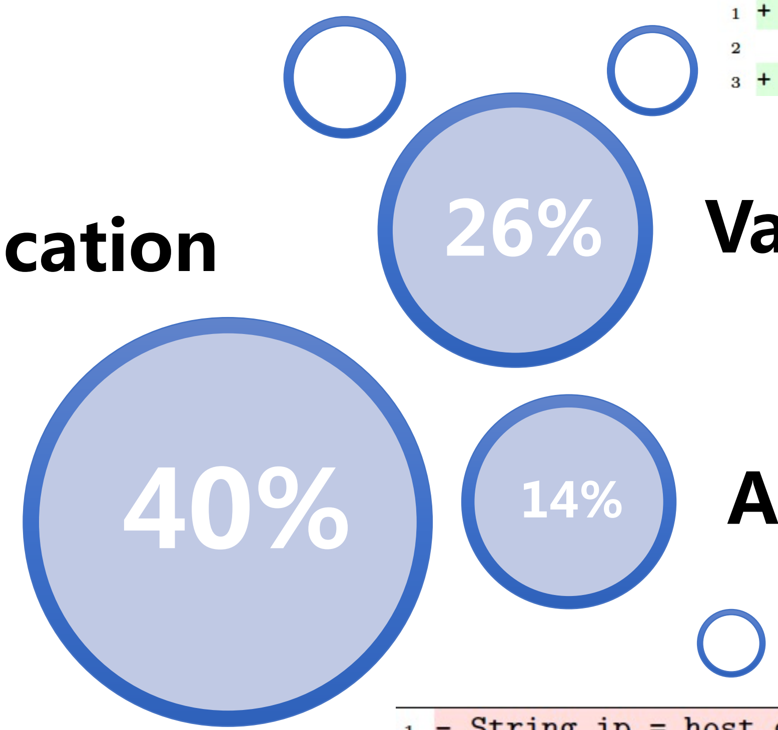
```
if (builder.length() > 1) {
  builder.deleteCharAt(builder.length() - 1;
}
```

```
if (@Para{sb, builder}.length() > 1) {
  @Para{sb, builder}.deleteCharAt(@Para{sb, builder}.length()-1);
}
```

Generic Patch Template

15 million commits
30 million files

Patch Category



API Modification

```
  multipleSource.setParams(
    MultiSourceConvertUtil.buildReqParams(
-     itemSku.getItemId().getValue(),
+             itemSku.getConfigId(),
              itemSku.getSkuId(),
              itemSku.getSellerId(),
+             itemSku.getGpuId()},
              multipleSource.getPageSize(),
              multipleSource.getPageIndex()));
```

Validation Check

```
+ if (accountStatus != null && accountStatus.length > 0) {
    query.addCondition(new In(STR, accountStatus));
+ }
```

API Wrap

```
- String ip = host.getIp();
- String url ="http://".concat(ip).concat(flowHost);
- HttpResponse<String> response = httpRequest.asString();
- ErrorUtil.checkError(httpRequest, response, TREND, start);
- bodys.add(response.getBody());
+ String url = UrlUtil.getUrl(headHost, flowHost);
+ UrlUtil.requestAndCheckThenFillResult(httpRequest, bodys, TREND, sta
```

Results

EFFECTIVENESS

22%

False positive rate is 22% in patch discovery and it is supposed to be gradually reduced by feedback on discovered patch and contribution of new patch

USER STUDY

10/12

The majority (10/12) of the interviewed developers acknowledged the value of the patches, and all of them would like to see Precfix adopted in practice

EFFICIENCY

5 Hours

Offline patch discovery costs 5 hours (extracting pairs, clustering, and extracting templates consumes 22, 270, and 5 min). Online patch recommendation is made within milliseconds

DEPLOYMENT

1 Year

Precfix has been deployed in Alibaba for about one year so far. Every week, it recommends about 400 patches to developers on average, and receives about two to three false positive reports