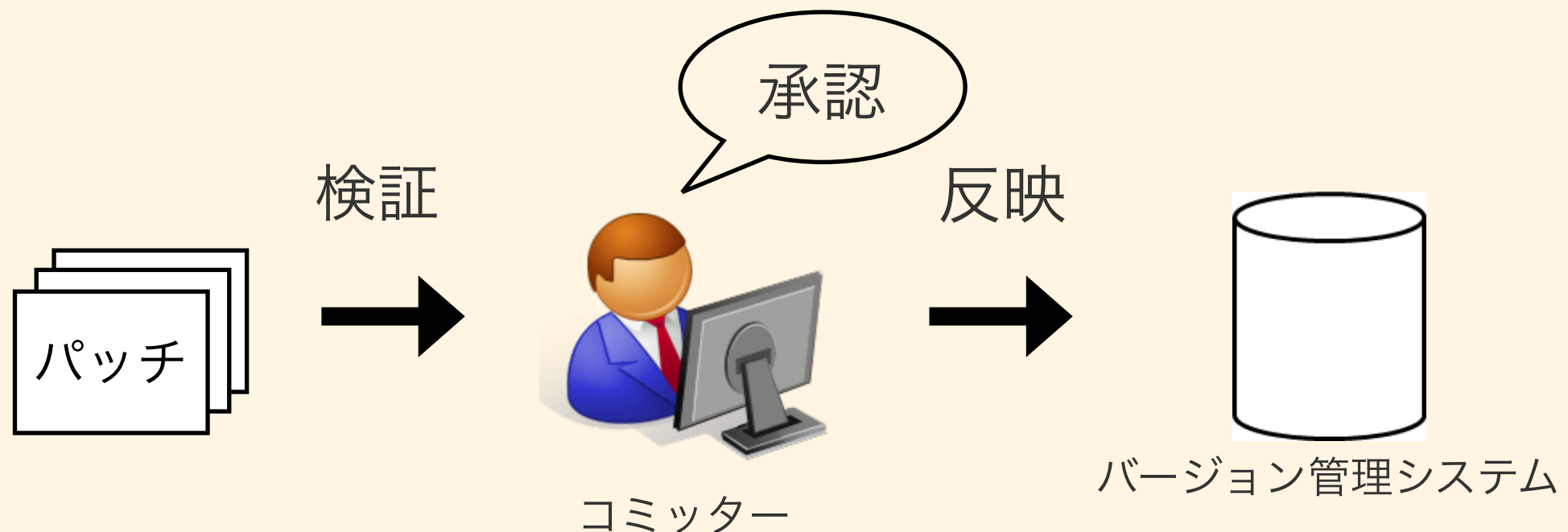


昇格時期に基づいた コミッター候補者予測 (仮)

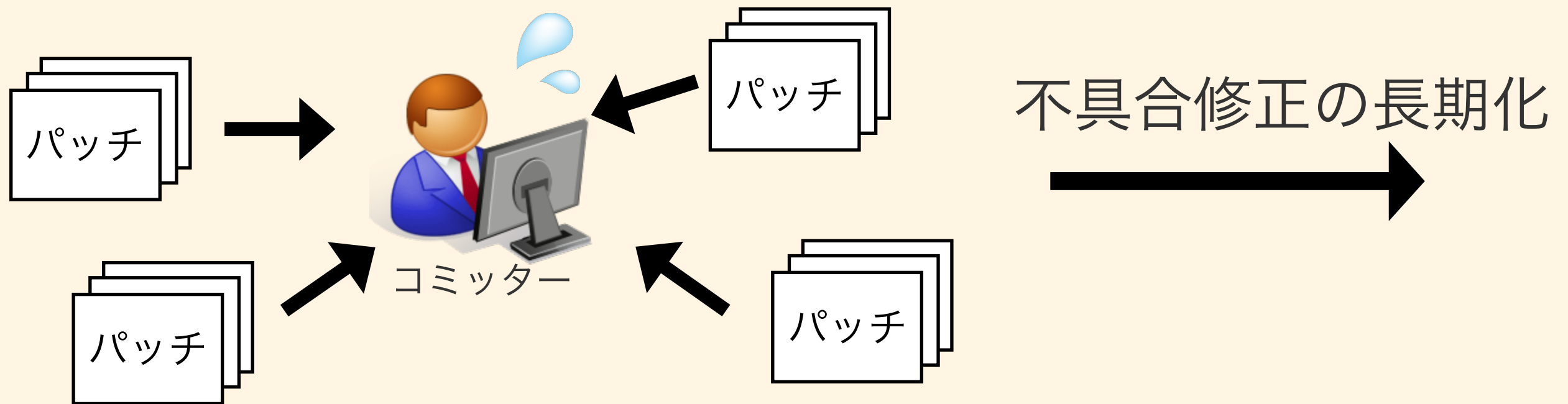
背景

- オープンソースソフトウェア(OSS)が社会に普及するにつれてプロジェクトへの不具合報告が増加している
- コミッター（インテグレータ）と呼ばれる開発者がパッチの検証・プロジェクトへの反映を行う



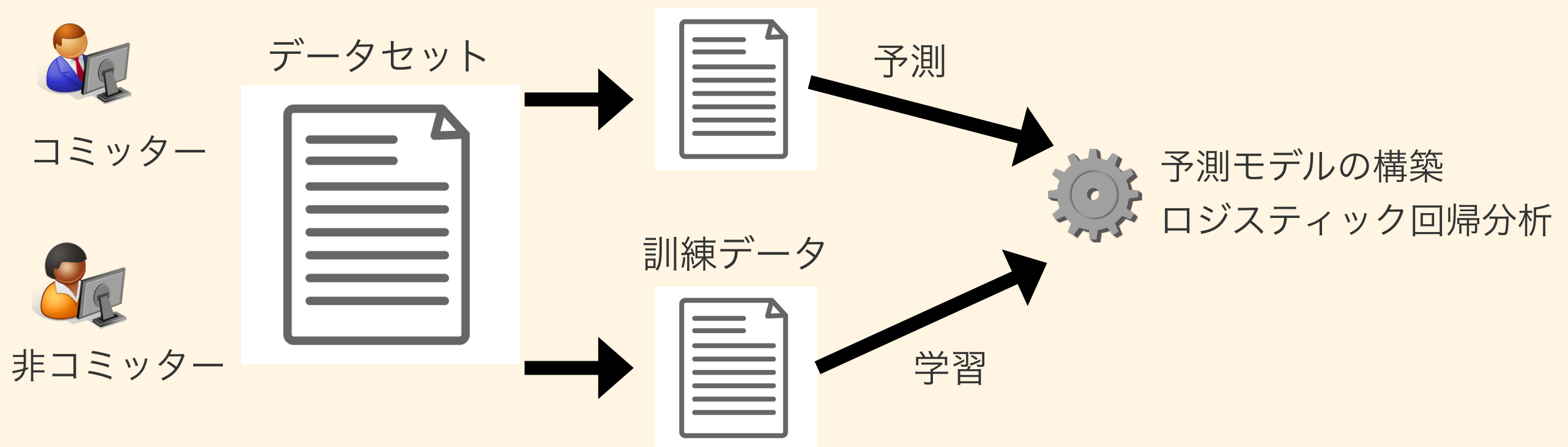
コミッターの現状

- コミッターは開発者の中でもごく一部であり，一人当たりの作業負担が集中している
- パッチの検証・反映が滞り，**不具合修正が長期化**している



既存研究

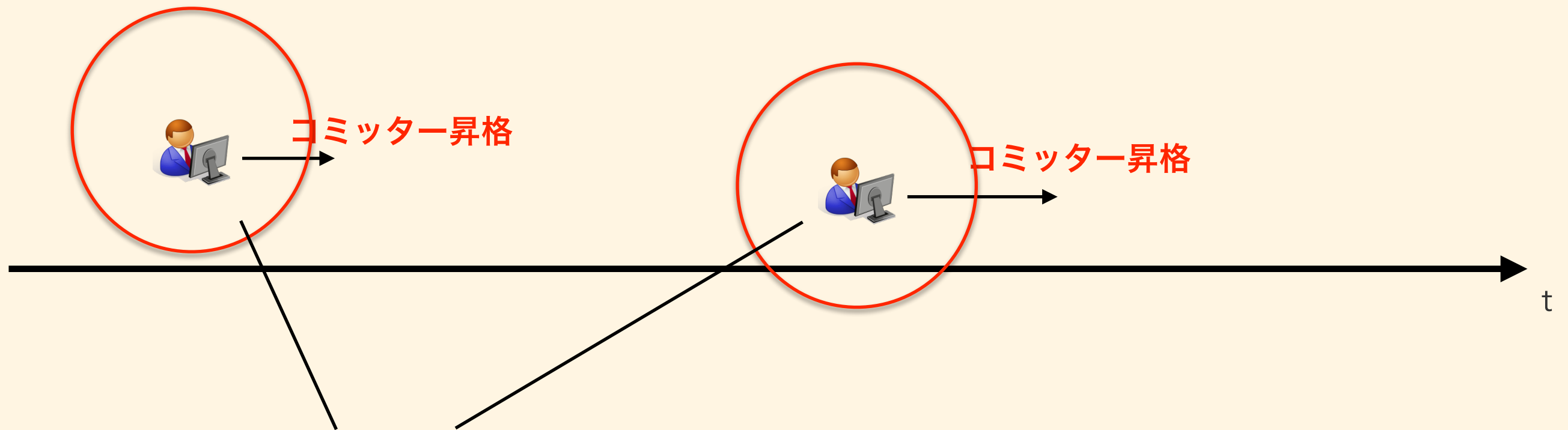
- コミッターと非コミッターの活動量メトリクスを元にコミッター候補者を予測するモデルの構築をしている^[1]



[1]Gharehyazie, M., Posnett, D., Vasilescu, B., & Filkov, V. (2015). Developer initiation and social interactions in OSS : a case study of the Apache software foundation. *Empirical Software Engineering*, 20(5), 1318-1353. DOI: 10.1007/s10664-014-9332-x, 10.1007/s10664-014-9332-x

問題点

- コミッターの昇格時期を考慮せず、全期間で予測している
- 時期によってプロジェクトの状況が異なっている可能性はある



昇格時期に着目して予測した方が高くなる（はず）

目的とアプローチ

- 目的
 - コミッター予測の精度の向上
- アプローチ
 - 昇格時期のプロジェクト状況を考慮し、予測を行う

やっていること

- 既存研究でコミッターと分類された開発者を目視した結果、すでにコミッターと思われる開発者が多かった
- 現在の方法（コミット以前に貢献を行った開発者をコミッターとする）以外でコミッターを定義したほうが良いと考えられる

やっていること

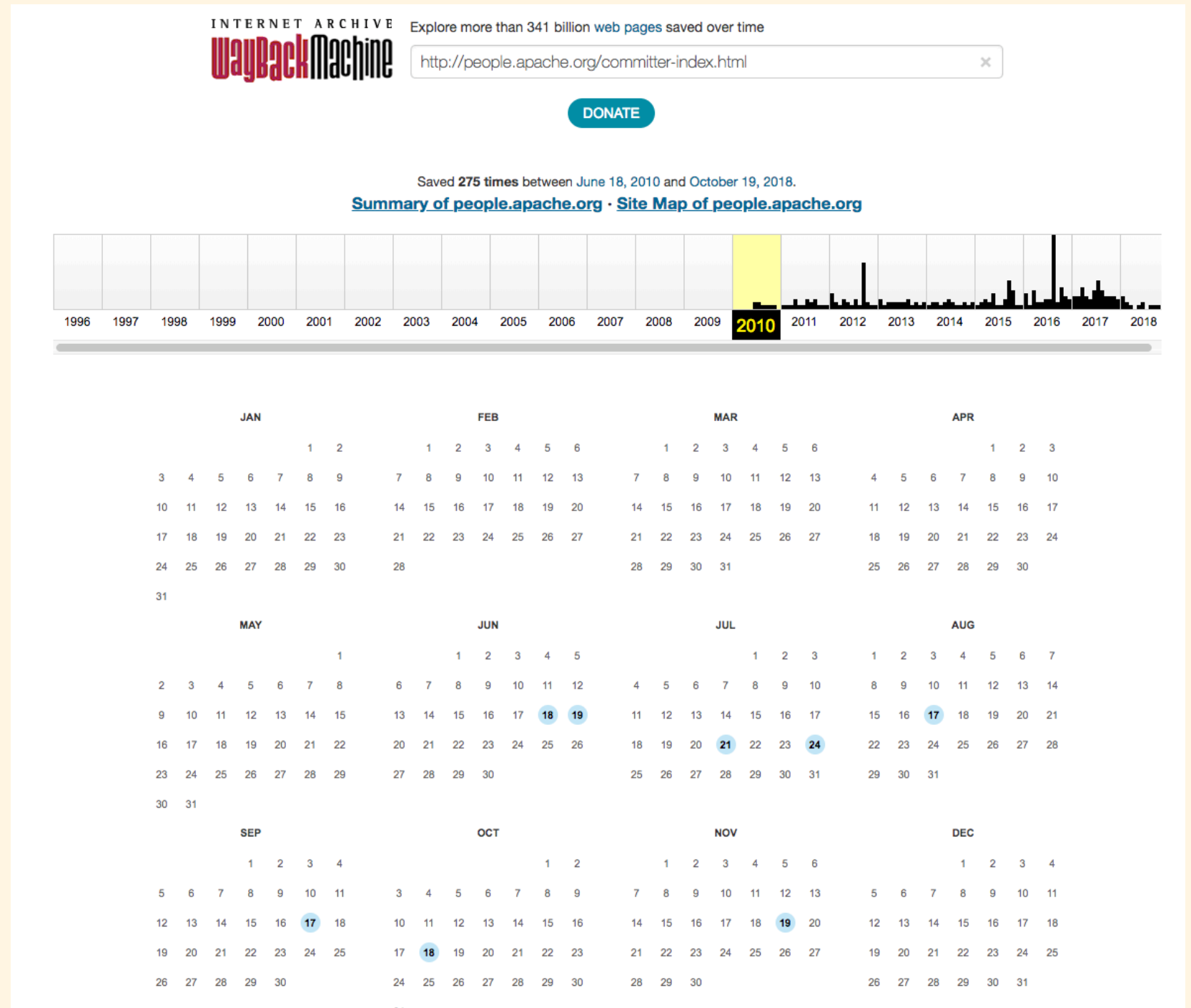
- asf committer list に載った開発者をコミッターとし

て定義

 ASF Committers by Id		
This page lists all known committers by login id. For each entry, it shows the full name and any LDAP groups / SVN groups to which they belong. This information is derived from the JSON files which in turn are derived from LDAP and the SVN authorization file by Whimsy. Entries in bold are ASF members. Committers may provide homepage URLs in LDAP. Login to https://id.apache.org/ and populate the "Homepage URL:" field. Remember to provide your password (near the bottom of the page) before pressing submit. Any such entries are shown as links in the Name column. There is also a list of Persons with signed CLAs but who are not (yet) committers. This is useful for checking if a CLA has been received and recorded in the iclas.txt file which is maintained by the secretary. If the committers has a PGP key then it is linked from the SVN Id column Please note that some existing ids include special characters such as '-' and '_'. These characters are no longer permitted. Ids must be lower-case alphanumeric (ASCII) only, at least 3 characters and must start with a letter. Last updated at: 2018-10-30 00:10 UTC		
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z		
SVN id	Name	Group membership
a_budroni	Alessandro Budroni	incubator
a_horuzhenko	Artyom Horuzhenko	openmeetings-pmc , openmeetings
aa	Andrew Ahn	atlas-pmc , atlas , incubator
aadamchik	Andrus Adamchik	apsite , cayenne-pmc , cayenne , incubator-pmc , incubator , member
aadomowski	Aleksander Adamowski	directory
aadrian	Adrian Antal	
aaf	Alexei Fedotov	incubator , openmeetings-pmc , openmeetings
aahepp	Aaron Hepp	openmeetings-pmc , openmeetings
aajisaka	Akira Ajisaka	hadoop-pmc , hadoop , yetus
aalmiray	Andres Almiray	groovy-pmc , groovy , incubator , tamaya
aanderson	Aaron Anderson	ode
aaron	Aaron Bannert	apr-pmc , apr , apsite , httpd , incubator-pmc , incubator , infrastructure , member
aaronfeng	Aaron Feng	incubator , madlib-pmc , madlib , pmc-chairs
aasanaliev	Askhat Asanaliev	incubator , usergrid-pmc
aavenoso	Andrew Avenoso	incubator
ab	Andrzej Bialecki	apsite , gora-pmc , gora , incubator-pmc , incubator , lucene-pmc , lucene , member
ababiichuk	Andriy Babiichuk	ambari-pmc , ambari , incubator
abafna	Abhishek Bafna	oozie
abai	Amy Bai	hawq-pmc , hawq , incubator
abaker	Anthony Baker	geode-pmc , geode , incubator , member
abaranchuk	Artem Stepanovich Baranchuk	ambari
abarve	Amey Barve	geode , incubator

やっていること

- スナップショットを保存しているサイトを用いて差分を抽出する



やりたいこと

- プロジェクトの状況を測るメトリクスを作成
 - するために活動を目視して使えそうなものを考えてみる
- 別アルゴリズムでの予測
 - ロジスティック回帰は今回の予測に向いていないと考えられる
-

プロジェクトの状況を測る メトリクスを作成

- ~~~~~

別アルゴリズムでの予測

コミッタ昇格(既存手法)

	1年	2年	3年	4年	5年	6年	7年	8年	9年	10	11	12	13
Ant	4	8	4	4	4	2	0	1	3	1	0	0	
Axis2_	0	4	7	6	4	0	0	0					
Log4j	3	1	2	1	1	0	1	0	0	0	1		
Lucen	6	0	5	3	4	4	2	4	2	2			
Pluto	4	1	1	2	2	3	0	0					
Solr	5	1	3	2	3	0							

ant 1年目

p	l	pl	neidevs
6	151	157	6
85	78	163	2
146	170	316	7
127	137	264	1

ant 1年目 sum patch

1,4人目が多い

sumpatch_p0	sumpatch_p1	sumpatch_p2	sumpatch_p3	sumpatch_p4	sumpatch_p5	sumpatch_p6	sumpatch_p12	sumpatch_p24
0	0	2	3	5	9	11	13	17
0	0	0	0	0	0	0	0	2
0	0	0	0	0	0	0	0	0
0	0	0	11	14	17	19	19	19

sumpatchinline_p0	sumpatchinline_p1	sumpatchinline_p2	sumpatchinline_p3	sumpatchinline_p4	sumpatchinline_p5	sumpatchinline_p6	sumpatchinline_p12	sumpatchinline_p24
0	0	2	3	5	5	5	6	8
0	0	0	0	0	0	0	0	2
0	0	0	0	0	0	0	0	0
0	0	0	11	11	11	12	12	12

sumpatchattach_p0	sumpatchattach_p1	sumpatchattach_p2	sumpatchattach_p3	sumpatchattach_p4	sumpatchattach_p5	sumpatchattach_p6	sumpatchattach_p12	sumpatchattach_p24
0	0	0	0	0	4	6	7	9
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	3	6	7	7	7

ant 1年目 neighbor

1,3人目が多い 4人目が少ない

neinum	neinum_p1	neinum_p2	neinum_p3	neinum_p4	neinum_p5	neinum_p6	neinum_p12	neinum_p24
2	9	12	14	26	48	86	148	214
0	2	6	18	31	35	42	78	173
0	23	63	74	86	94	108	191	289
1	1	4	5	6	8	11	20	26

neidevs_p0	neidevs_p1	neidevs_p2	neidevs_p3	neidevs_p4	neidevs_p5	neidevs_p6	neidevs_p12	neidevs_p24
2	4	5	5	5	6	6	6	7
0	0	1	2	2	2	3	3	4
0	4	5	5	5	5	7	10	11
0	0	1	1	1	1	2	2	2

ant 1年目message, thread

1,3人目が多い

nmsg_p1	nmsg_p2	nmsg_p3	nmsg_p4	nmsg_p5	nmsg_p6
14	18	23	41	99	270
2	12	45	135	152	183
54	189	227	264	304	381
2	7	9	11	15	19

nthr_p1	nthr_p2	nthr_p3	nthr_p4	nthr_p5	nthr_p6
0	2	3	5	7	32
2	4	11	54	195	255
0	0	0	0	3	5
5	19	57	65	82	110

2,4人目が多い

別アルゴリズムでの予測

- コミッターごとに活動内容が異なる（パッチ投稿が多かったりメーリスなどでサポートをしたり）ため

3人目をテストデータ, 4人目を訓練データとした場合予測できない可能性がある

予定

- 10月29日~11月02日 コミッタリストの作成&目視
- 11月05日~11月09日 実験
- 11月12日~11月16日
- 11月19日~11月23日 論文執筆
- 11月26日~11月30日
- 12月03日~12月07日
- 12月10日~12月14日 論文（日本語版）完成