

Identifying and Quantifying Architectural Debt

Lu Xiao, Yuanfang Cai, Rick Kazman, Ran Mo, Qiong Feng

60191055 山崎大輝

はじめに

欠陥のある建築関係はファイル間で欠陥を伝播し、時間の経過と共にまるで債務のように高いメンテナンス費を累積する。

これをArchitectural Debtと呼び、こういったファイル郡をモデル化していく。

7つの大規模オープンソースプロジェクトに対しアプローチを行い、

Architectural Debtへの関心を払うことでプロジェクト全体のメンテナンスのかなりの部分が消費されていることを示す。

Technical Debt(TD)について

近い将来の目標を達成するためにコーディング活動で取られたショートカットの長期的な影響を記述するメタファー

要するに手抜きした代償

こういった欠陥は時間の経過と共にメンテナンスコストを増加させていく。

背景

DSMを用いてDRSpaceを表現している。

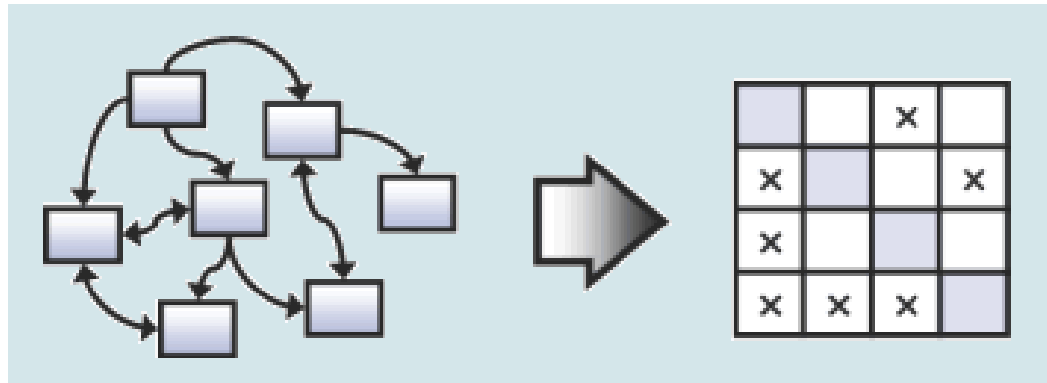
DRSpace・・・Design Rule Space

システムのソースファイルのサブセットであり、これらのファイル間の何らかの関係(依存関係)

DSM・・・ Design Structure Matrix

DSMについて

ソフトウェアのモジュール構成を展開して、それぞれの依存関係をマトリクス表示する。



DSMについて

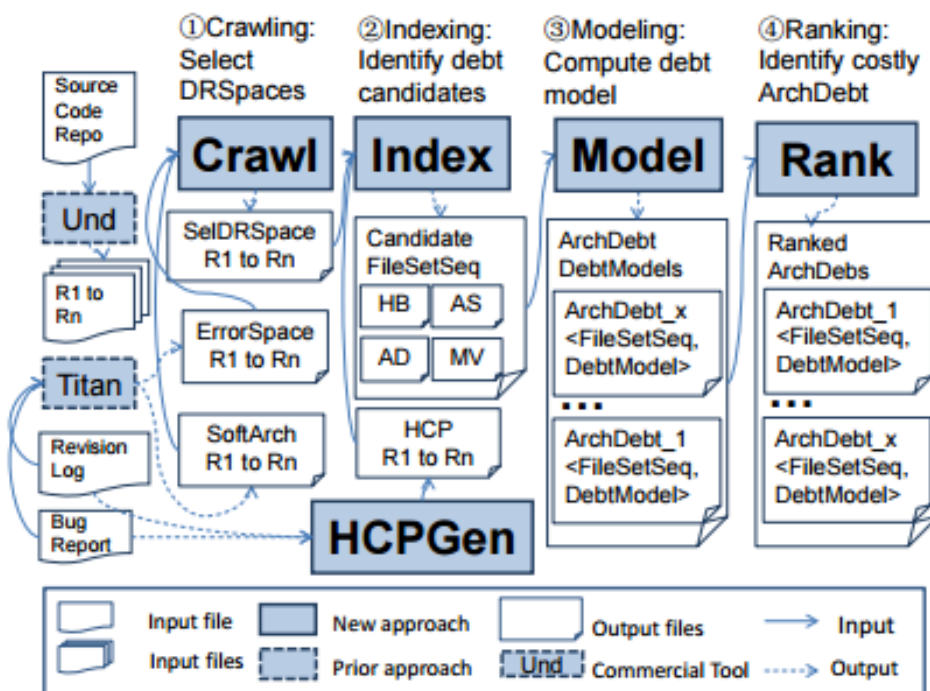
セル[6,2]には“Implement”が含まれている。つまり、6行目のファイル(CassandraServer)は2行目のCassandraにインターフェイスを実装している。

	1	2	3	4	5	6
1 ColumnParent	(1)	,100%	,50%	,41%	,50%	,100%
2 Cassandra	dp,	(2)	,48%			
3 CliClient	dp,	dp,	(3)			
4 Column*Reader	dp,	dp,	(4)			
5 ThriftValidation	dp,			(5)		
6 CassandraServer	dp,	Implement,			dp,	(6)

セル[2,6]には「48%」が含まれている。つまり、カサンドラが変更されると、カサンドラサーバーが変更される確率は48%である。

ArchDebtの識別

FileSet-Sequenceを特定し、FileSet Sequenceが発生したコストに基づいて「金利」を取得するためのDebt Modelを構築します



ArchDebtの識別

- 1) クロール: Webページをクロールして収集するのと同様に、tEacherのそれぞれからDR空間のサブセットを1からRに集めます。
- 2) インデックス: このステップは、特定のファイルを識別(索引付け)する
- 3) モデリング: FileSetrの各シーケンスによって発生するメンテナンスコストを測定する。
- 4) ランキング: ArchDebtの重大度は、プロジェクトの進化履歴に蓄積されている保守コストの量に応じてランク付けします。

評価

RQ: ArchDebtsで特定されたファイルグループが大量の保守コストを生成し、増加させるかどうか。つまり、彼らは本当に重大な借金か？

特定されたファイルグループが全体的な保守作業のうちのほんの一部を消費する場合、注意を払う必要はない。

同様に、特定されたファイルグループがシステムそのものの大部分を占めている場合、それらが大半の保守作業をも消費するかどうかは驚くべきではない

評価科目

7つのApacheオープンソースプロジェクトを選択した。

Camel - エンタープライズ統合パターン。

Cassandra - 分散型DBMS

CXF - Webサービスフレームワーク

Hadoop - 信頼性の高いスケーラブルな分散コンピューティングのフレームワーク

HBase - Hadoopは分散型のスケーラブルな大容量データストア。

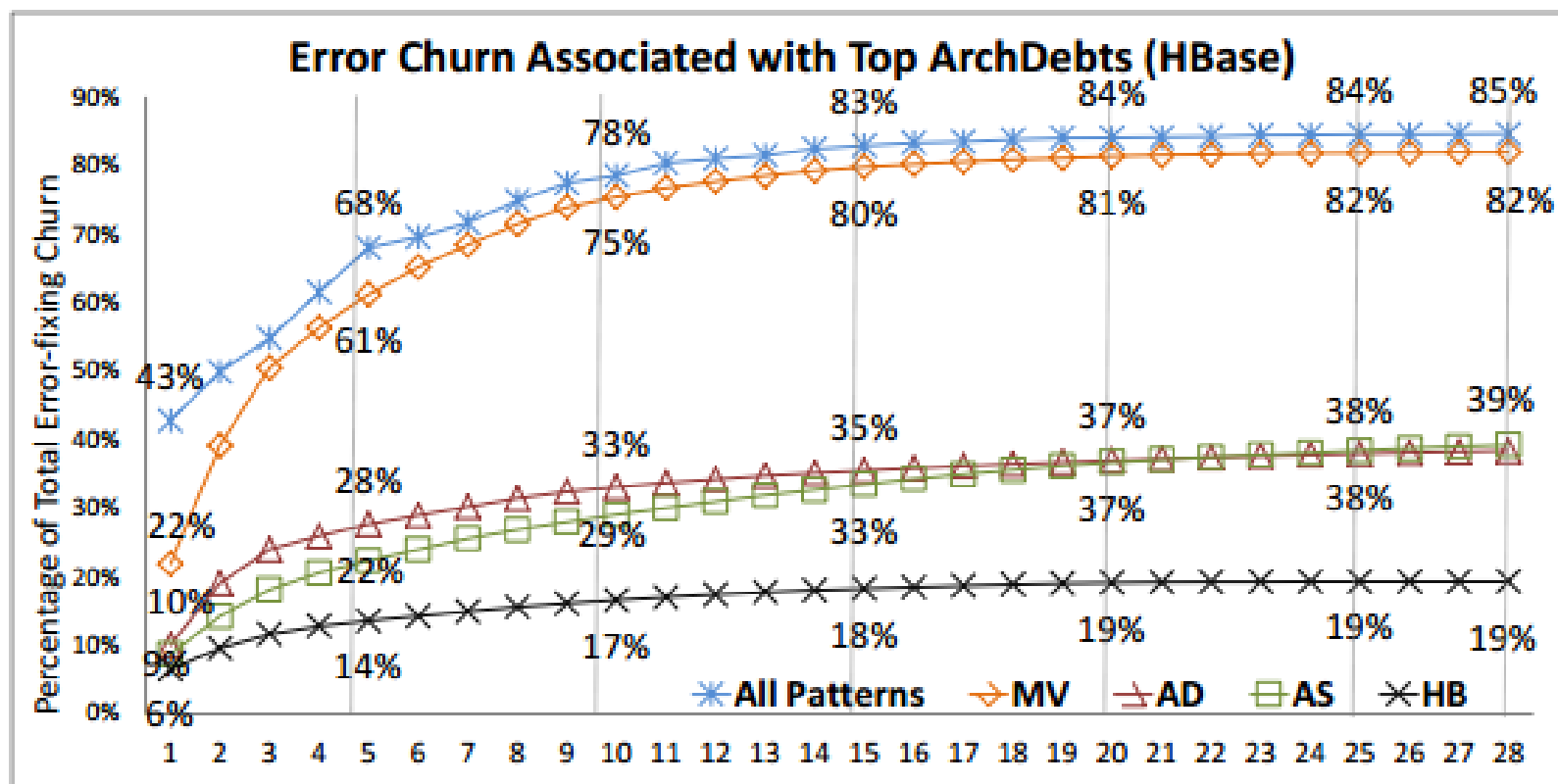
PDFBox - PDFドキュメントを扱うためのライブラリ。

WicketはコンポーネントベースのWebアプリケーションフレームワーク。

各プロジェクトの概要

Subject	Length of history (#Mon)	#R	#Cmt	#lss	#Files
Camel	7/2008 to 7/2014 (72)	12	14563	2790	1838 to 9866
Cassandra	9/2009 to 11/2014 (62)	10	14673	4731	311 to 1337
CXF	12/2007 to 5/2014 (77)	13	8937	3854	2861 to 5509
Hadoop	8/2009 to 8/2014 (60)	9	8253	5443	1307 to 5488
HBase	12/2009 to 5/2014 (53)	9	6718	6280	560 to 2055
PDFBox	8/2009 to 9/2014 (62)	12	2005	1857	447 to 791
Wicket	6/2007 to 1/2015 (92)	15	8309	3557	1879 to 3081

調查結果



調査結果

- ①建築負債は、プロジェクト全体の維持努力のかなりの割合(85%)を消費する。
- ②建築上の負債の中では、メンテナンスが大部分を占めている。
- ③Modularity Violation負債は最も一般的で高価な負債である。

まとめ

評価した結果、全体的なメンテナンスのかなりの部分(51%～85%)が Architectural Debtによって消費されたことが示された。

→これらの債務を早期に発見し、リファクタリングによってそれらを払うことができれば、プロジェクトが相当量の維持費を節約できることを示唆している