

Two's Company, Three's a Crowd: A Case Study of Crowdsourcing Software Development

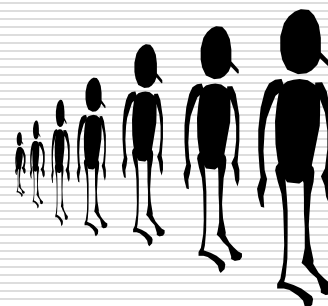
紹介者: 角田 雅照 (近畿大学)

著者:

Klaas-Jan Stol, and Brian Fitzgerald (Lero,
University of Limerick, Ireland)

□ **クラウドソーシング**：インターネット上で不特定多数の人を募り，仕事を依頼すること.

- 画像のタグ付けなど，一般には**小さく独立した作業**が多い.
- Amazon Mechanical Turkなど.



□ ソフトウェア開発においてクラウドソーシングを活用する場合に，重要な点を明らかにする.

- ソフトウェア開発のタスクは**独立しておらず，複雑で巨大**であり，長い期間がかかる.

□ タイトルの“Two’s Company, Three’s a Crowd”は英語のことわざ.

- **文献調査**に基づき、ソフトウェア開発においてクラウドソーシングを利用する際に重要な点を6つ挙げる.
 - タスクの分割：開発作業は**独立しておらず複雑であるため**、容易でない.
 - 調整とコミュニケーション：人数が増えると、**コミュニケーションのオーバーヘッド**が増加する.
 - 計画とスケジューリング：**締切**をいかに守らせるか.
 - 品質保証：**質の低い**作業結果が紛れ込む可能性がある.
 - 知識・知的財産：知的財産が**流出**し、競争力が低下する危険性がある.
 - 動機付けと報酬：いかに動機付けをするか？ **プロジェクトの成功**には重要である. 仕事の価格の設定も重要である.



□ 6つの重要な点に関し半構造的インタビューを行った.

■ 半構造的インタビュー: ガイドに基づきインタビューする.

□ 会社: TechPlatform Inc. (TPI; 仮名)

□ クラウドソーシングプラットフォーム: のTopCoder (TC)

□ アプリケーション: Titan



□ 論文の貢献点: クラウドソーシングの定義, 6つの重要な点の提示, インタビューによる詳細な分析

□ 感想

■ 新しい話題に取り組んでいる.

■ インタビューという, 労力のかかる作業をしている.

■ 定量的な分析はほとんどしていない.

□ 論文中の表は単なる集計結果である.

Does Latitude Hurt while Longitude Kills?

Geographical and Temporal Separation
in a Large Scale Software Development Project

Patrick Wagstrom

IBM Watson Group

Subhajit Datta

Singapore University of Technology and Design

ICSE勉強会 '14 資料

紹介担当者: 中川尊雄(奈良先端大)

目的と手法

• 目的

- 分散開発の既知とされる課題の真偽をデータ分析で確かめる

緯度が痛み、経度が殺す

- “Latitude Hurts, Longitude Kills”

- 単純な距離もさることながら、時差はいつそう開発者間のインタラクションを妨げる

• 手法

- IBMの開発チームにおけるWork Item Trackerの履歴を分析
 - 各々のWork Itemに対するコメントの応酬が分析対象
 - 65222コメント, 13399ワークアイテム, 17タイムゾーン

加工元画像: © MacDaddy /Vecteezy.com.



ひとつめの分析

- 距離・時差がコメントの応答時間に与える影響
- 重回帰分析によって以下の結果を導出

要素	係数	p-value
log(previous delay)	0.1445	<0.0001
log(repetition)	-0.1506	<0.0001
log(geographic distance)	-0.0919	<0.0001
temporal offset	0.1352	<0.0001

- 距離が離れるほど応答速度が速くなる
- 時差があるほど応答速度が遅くなる

ふたつめの分析

- 距離・時差が将来のコメント応酬確率に与える影響
- ERGMs(Exponential Random Graph Models)を利用
 - 非常に大雑把に言うとロジスティック回帰のグラフ版

要素	係数	p-value
reciprocity	6.1465	<0.0001
is_rational	0.5718	<0.0001
log(geographic distance)	-0.0744	0.0715
temporal offset	0.0372	<0.0001

- 距離があるほど将来のコメント応酬が起こりにくい
- 時差があるほど将来のコメント応酬が起こりやすい

所感

- IMやチャットとは異なる, そもそもの応答時間が長いツールを取り上げたことは不思議でもあるが, 面白い.
- データ分析が丁寧
 - コメント著者の出自(Rational Employee, IBMer, Non-IBMer)ごとに仕事場が異なるため, 応答速度の偏りを事前に確認
 - 回帰分析の際に, 変数の少ない組み合わせも試し, 詳細に記載
- Slideshareに発表資料があったのだが, Maybeが散見され, やはり結果がクリアに出なかったことを気にしている模様