

Twitter 発言の分析による Web サービス障害の影響調査

プロジェクトマネジメントコース 矢吹研究室 1442012 岩瀬翔

1. 序論

複数のメンバが同時に開発を行うソフトウェア開発プロジェクトにおいて、「GitHub」のような Web サービスが使われることがある。

Web サービスの停止は、それを利用しているプロジェクトに大きな影響を与えると思われる。実際、2016 年 1 月 28 日の GitHub の停止時には、そのせいで仕事が進められなくなったというようなツイートが Twitter 上で複数観測された [1]。

2. 目的

Twitter の発言を収集するためのツールを開発し、それを用いてソフトウェア開発で利用される Web サービスの停止が開発に与える影響を調査する。

3. 手法

2016 年に発生した GitHub の障害発生に関するツイートをデータとして収集する。Twitter の API には、1 週間以上前のツイートは取得できない制限がある [2]。そこで、制限無く検索できるブラウザの Twitter を利用する。この検索結果は画面を最下部にスクロールすることで古いものが読み込まれていく。検索する日付は GitHub を継続的に監視している「GitHub Status」を参照する。

データを収集するツールを開発するために 2 つのプログラムを作成する。1 つ目では、ブラウザの自動操作ができるライブラリである「Selenium WebDriver」を使用し、画面スクロール後、HTML ファイルを保存する。2 つ目では、Python のライブラリである「BeautifulSoup4」でデータを抽出する。

4. 結果

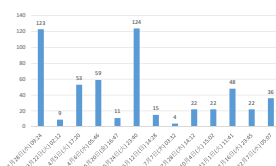


図 1 サービス停止から復旧までの間隔

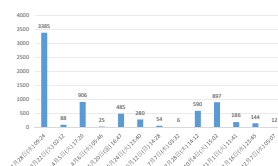


図 2 サービス停止中に投稿されたツイートの数

2016 年に GitHub で発生した 13 回のサービス停

止を調査対象とする。各障害のサービス停止から復旧までの間隔を分単位でグラフにしたものが図 1 である。そして、各障害のサービス停止中に投稿されたツイート数をグラフにしたものが図 2 である。

5. 考察

結果から、サービス停止時の曜日や時間帯によってツイート数に差はあるが、調査した全ての障害で反応が観測された。特に多かったのは平日の日中であったが、土日祝日や深夜でもツイートが観測されていることから、その影響が出ていることがわかる。このようなことから、Web サービスを使用する場合は、サービスが停止するリスクを考慮する必要があると考える。

また、GitHub Status に記されているアナウンスよりも、平均約 7 分早くサービス停止に関するツイートが観測されていた。これは、サービスの状態について運営元が発表している情報が必ずしも正しくはないことを示唆している。

6. 結論

Twitter のブラウザでの検索結果を保存するツールを開発し、それを用いて GitHub や Slack などの Web サービスの停止に対する開発者の反応を調査した。その結果、日中はもちろん深夜でもサービス停止の影響は大きいこと、サービス運営元による停止時間についての発表は実際のそれとはずれていることがわかった。このように Web サービスの障害とその影響を調査することが、それを利用するソフトウェア開発のマネジメントにおいて有用な知見となることが期待される。

参考文献

- [1] 岩城俊介, @ IT. Github がダウン、「ぜんぶのせいだ」のような惨状 今後の課題も. <http://www.atmarkit.co.jp/ait/articles/1601/28/news126.html> (2016.06.29 閲覧)。
- [2] 鳥海不二夫. Twitter 上のビッグデータ収集と分析. 組織科学, Vol. 48, No. 4, pp. 47–59, 2015.