

WIDE Technical-Report in 2017

WIDEクラウドWG 2016年度活
動報告
wide-tr-cloud-report-summary-2016-00.pdf



WIDE Project : <http://www.wide.ad.jp/>

*If you have any comments on WIDE documents, please contact to
board@wide.ad.jp*

Title: WIDE クラウド WG 2016 年度活動報告
Author(s): WIDE クラウドワーキンググループ (cloud-wg@wide.ad.jp)
Date: 2017-1-6

WIDE クラウド WG 2016 年度活動報告

WIDE クラウドワーキンググループ

WIDE クラウドワーキンググループは、今後のクラウド技術の研究開発を推進するために 2010 年 1 月に設立された。複数の WIDE 組織間に渡って運用される広域連邦型クラウドシステムである WIDE クラウドシステムの運用と、それをを用いた研究開発を行っている。

2016 年度は、システムログ解析技術の研究開発を実施した。

因果推定によるログの関係情報抽出

昨年度から引き続き、システムログイベント間の因果関係を PC アルゴリズムを用いて推定し、障害発生の原因推定、事前予測に活用するための技術を研究している。システムログを自動的にテンプレート化し、各テンプレートをイベントとみなすことによりイベントの因果関係を自動的に抽出する。実験では、SINET4 のネットワーク機器のログから BGP の再接続処理に起因して別のネットワーク機器で発生したインターフェースエラーを発見することに成功している。本研究は、その成果の一部を ACM SIGCOMM CoNEXT の学生ワークショップにて発表している [1]。

新しいテンプレート化手法

システムログメッセージを分類する際、固定部と変数部を検出するテンプレート化を用いることが一般的に行われている。本研究では、新しいテンプレート化手法として、メッセージの各単語の長さを元にした手法を検討した。過去の多くのテンプレート化手法が、二段解析による負荷の高い手法や、文字種別の比率などに基づく複雑な手法を用いていたのに対し、単語長のみで一定の品質のテンプレートが生成できることを確認した。本結果はテクニカルレポートとして公開している [2]。

共有ワードグラフでの状態推定

システムログメッセージを解析する場合、テンプレート化を実行してテンプレートごとの出現頻度などを調査する手法が広く採用されている。しかし、そもそものメッセージが自由記述の文章であることや、近年の複雑なソフトウェア構成事情から、正確なテンプレートを生成するのは困難かつ時間のかかる作業となりつつある。本研究では、システムログメッセージからシステムの状態を推定する際、テンプレート化作業を実施しない方法を模索する。

システムログメッセージを固定行数ごとに区切り、それぞれの行間で共有される単語を抽出する。各行を一つのグラフノードとみなし、共有単語を持つノードを接続していくことにより、行数と等しいノード数を持つ無向グラフが生成される。類似したメッセージが出力される場合、生成される無向グラフも類似したものとなる。このグラフを共有ワードグラフと呼称し、その時間のシステム状態を表現するとみなすこととした。このグラフ間の距離を Sorted graph histogram[3] を用いて観測、分類することにより、通常状態（頻繁に観測される状態）から非通常状態（稀に観測される状態）を発見できると考えている。

まとめ

より詳しい報告は別途配布される詳細報告書を参照して欲しい。大規模化の進むネットワークサービスを安定して運用するためにはより良いシステム状態推定技術が必要である。引き続き研究開発を進めていく。

参考文献

- [1] Satoru Kobayashi, Kensuke Fukuda, and Hiroshi Esaki. Causation mining in network logs. In *ACM SIGCOMM CoNEXT 2016 Student Workshop*, 2016.
- [2] Keiichi Shima. Length matters: Clustering system log messages using length of words. *CoRR*, Vol. abs/1611.03213, , 2016.
- [3] Apostolos N. Papadopoulos and Yannis Manolopoulos. Structure-Based Similarity Search with Graph Histograms. In *Proceedings of the tenth International Workshop on Database and Expert Systems Applications (DEXA'99)*, pp. 174–178, 1999.