

ストリームデータモニタリングにおける確認タイミングの 判断支援に関する予備的検討

Preliminary Study on Support of Determining Timing to Monitor Stream Data

吉田和人* 高間康史

Kazuhito Yoshida, Yasufumi Takama

首都大学東京システムデザイン学部

Faculty of System Design, Tokyo Metropolitan University

Abstract: This paper reports the preliminary study on the development of monitoring support system for stream data. It is supposed that stream data such as online news has to be monitored during break of user's primary job. If a user check a stream data at wrong time, the efficiency of his/her primary job would go down. In order to help a user to monitor stream data, we are developing a system that gives a user a clue for determining the timing to monitor with using a dynamic bar chart. This paper reports the result of preliminary experiment, in which the effect of color of bars and individual differences on the timing decision is investigated.

1. はじめに

本稿では、オンラインニュース等のストリームデータを確認するタイミングの判断を支援するシステムの構築に向けて、棒グラフ形式のメータを利用したユーザ実験を行った結果について報告し、棒グラフの色や個人差の影響について考察する。

近年、ストリームデータの量は膨大になっており、ヤフー株式会社の提供するニュースポータルサイトYahoo!ニュース¹から提供されている「主要」カテゴリのニュースだけでも一日当たり70-100件程度配信されている。また、Twitter²では一日当たり5億件のツイートが発生している³。

これらのストリームデータは有益な情報も多く含んでおり、定期的に観覧しているユーザは多数存在するが、個人が全ての情報を常時確認することは通常不可能である。従って、情報のモニタリング支援を行う必要があると考える。

モニタリング支援に対するアプローチの一つとしてニュースキュレーションサイトが挙げられる。Gunosy⁴では、ユーザの趣向から自動的にニュースを選別し、提供している。これらのサービスは膨大なストリームデータの中から関心のある情報を効率的

に発見する作業を支援している。しかし、それらのニュースを確認するタイミングについては、朝刊や夕刊などの定時配信程度であり、個人及びその状況に基づく適切な確認タイミングなどは考慮されていない。

ストリームデータの確認は、休憩時間などの本務の合間に行われる作業とみなすことができる。したがって、モニタリングを頻繁に行えばその都度本務が中断されることになる。反対に、モニタリングの間隔を大きくすれば、ストリームデータの確認に時間を要する可能性があり、本務の中断時間が長くなる。一般に、作業に割り込みが発生すると、知的生産性が低下することが指摘されている。中断後に再開したタスクは中断されない場合の二倍時間がかかること[1]、中断されたタスクの40%は再開されないこと[2]などが指摘されている。従って、適切なタイミングでモニタリングを行うことは、効率的なストリームデータの確認のためだけではなく、円滑な本務遂行の上でも重要と考える。

本稿では、ストリームデータを確認すべきタイミングをユーザが適切に判断することを支援するシステムの構築を目的とする。本務を中断可能なタイミングを計算機が推定する割り込み可能性推定とは異

¹ <http://news.yahoo.co.jp/>

² <https://twitter.com/>

³ <https://about.twitter.com/ja/company>

⁴ <https://gunosy.com/>

*連絡先： 首都大学東京大学院システムデザイン研究科

〒191-0065 東京都日野市旭ヶ丘 6-6

E-mail: ytakama@sd.tmu.ac.jp

なり、ユーザ自身がタイミングを判断することを想定している。支援システムでは、ストリームデータの蓄積量を可視化により提示することを想定している。この時、ユーザが適切なタイミングで判断できるように、データ量と視覚的変数のマッピングをユーザごとに調整する必要があると考える。そのため、予備実験として、棒グラフ形式のメータを利用したユーザ実験を行った。その結果に基づき、棒グラフの色や個人差の影響について考察する。

2. 関連研究

2.1. 割り込み可能性推定

スマートフォン等の普及に伴い、人とのコンタクトをいつでも取ることができるようになった。しかし、仕事などの別の作業に集中しているときにこのようなコンタクトがあると作業を中断しなければならない。再び作業に戻る際に先ほどまで何をしていたか思い返す必要があるため、1節で述べたように作業効率の低下につながる。このような問題に対し、作業への割り込みをいつ行うと本務への影響が少なくなるかを推定する研究が行われている[1,2,3,4,5]。

ユーザの作業を推定するには、ユーザの状態を観測する必要があり、様々な手法が提案されている。卓ら[3]は推定可能な作業の汎用性を考慮してリストバンド型センサの3軸加速度データを用いてユーザの状態を取得している。谷ら[4]はセンサを身体につける煩わしさの観点から机にかかる圧力のデータを用いてユーザの状態を取得している。田中ら[5]は、検知の容易さ、PCを用いた作業との親和性の観点から、利用アプリケーションの切り替えデータを取得して推定に用いている。

2.2. 情報可視化システム

ストリームデータを可視化することによるモニタリング支援システムが研究されている[6,7,8]。

沼野ら[6]は定期的なオンラインニュースのモニタリング作業を支援するインタフェースを提案している。ニュース記事の文章クラスタリングに基づく話題の検出、追跡を行い、それらを新着記事、話題記事ごとの確認を行うリストモード、関心のある話題の新着記事数を確認できる続報記事確認モードの二つのモードを用いて可視化を行っている。続報記事確認モードでは続報記事数を黄色い四角形の数で可視化している。

奥村ら[7]は定期的な BBS (電子掲示板) のスレッ

ドのモニタリング作業を支援するインタフェースを提案している。キーワードベースの可視化を採用し、現在関心のある話題に関する投稿を追跡可能であるほか、新たな話題の発見も可能となっている。また、特定のスレッドから抽出したキーワード、複数スレッドから抽出したキーワードをそれぞれ別のビューで提示する、Overview + details を採用している。

黒澤ら[8]は OSS (オープンソースソフトウェア) の複数バグ管理システムから継続的に配信されるバグ更新情報のモニタリング作業を支援するインタフェースを提案している。報告されたバグはノードとして可視化され、ノードの大きさによりバグの修正に向けた進展や議論の進捗、色によりバグの修正状態を表現している。また、前回確認時から変化していない部分の縮小表示や軌跡の描画によって変化部分の確認を容易にすることで、効率的なモニタリングを支援している。

3. 予備実験用インタフェース

本稿では、モニタリングすべきタイミングを判断する手がかりをユーザに提示する汎用的な手段を検討する。テキストストリームデータの種類によらず必要な手がかりとして、モニタリングしていない間に到着したストリームデータの蓄積量を提示することを考える。本稿では、ストリームデータの蓄積量を棒グラフにより可視化する。データ量にマッピングすべき視覚的変数として、棒グラフの高さだけでなく色も併用することで、ユーザが視認しやすくなることが期待できるが、明度や色相など、視覚的変数として利用可能な色の属性は複数存在し、またユーザにとっても視認しやすさが異なる可能性がある。そこで、本稿では色やユーザ毎の特性の違いについて考察するために予備実験を行う。

予備実験に用いたインタフェースのスクリーンショットを図1に示す。4節に後述するように、予備実験は二種類行っているが、図1右の確認ボタンの有無の違いと、本務に相当するアプリケーションと同一の PC 上で動作するかの違いがある他は、両実験で用いたインタフェースは同様のものである。

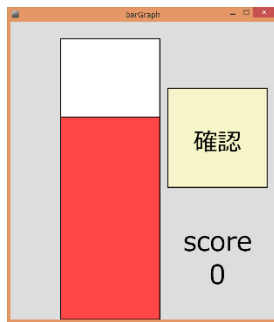


図 1. 実験に用いたインターフェース

図 1 の棒グラフはデータの蓄積量を可視化している。あらかじめデータ数の上限を設定しておき、上限の値で棒グラフの量が最大となるようにマッピングしている。インターフェース上の確認ボタンをクリックするか、Space キーまたは Enter キーを押すことで、データの確認を行ったとみなし、蓄積量を 0 に戻す。本稿での予備実験では、実際のストリームデータではなく人工的なデータの蓄積量を反映した。詳細については 4 節にて後述する。

本稿の実験で視覚的変数として用いている色のパターンは図 2 に示す 5 種類である。棒グラフの量の増加に従い、棒全体の色が変化する。

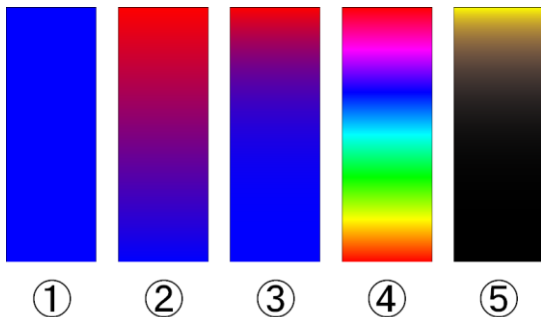


図 2. 実験で視覚的変数として用いた棒グラフの色パターン

①は青単色であり、色が変わらない場合のサンプルとして利用する。②は青から赤へ等比率で変化する。③は②と同様に青から赤への変化であるが、棒グラフ上方部で大きく変化するように調整している。④は色相環を一周するパターンである。⑤は明度の変化である。③と同様に、棒グラフ上方部で大きく変化するように調整している。②と③では、青と赤は信号など一般的に利用される色であるため利用した。⑤の明度は大きさを表現するのに適している[9]ことから利用した。また、今回の実験で行う

タスクにおいては、棒グラフの高さを判断の手がかりにするだけでなく、特定の色を手がかりとする可能性も存在するため、違いを認識しやすい色相を④で利用している。

本稿での実験では、実験協力者にデータの最大値になるべく近い所でリセットしてもらうための動機づけとして、スコアを導入している。スコアはリセットを行うたびに加算される。リセット時の棒グラフの高さを、最大値に対する割合 (%) で求め、その値を二乗したものをスコアとする。例として、棒グラフの高さが 50% の時にリセットを行った場合 $50^2 = 2500$ 点が加算される。ただし、リセット前に最大値を超えた場合は強制的にリセットされ、スコアとして -10000 が加算される。

4. 予備実験

4.1. 実験概要

20 代の工学系大学生、大学院生を対象に、前節で述べたインターフェースを利用した実験を実験 (1)、(2) の 2 回に分けて行った。実験 (2) では実験 (1) で得られた知見や反省に基づきタスクや設定などを変更している。

本実験の目的は、確認タイミングの判断支援のために用いる棒グラフのような、直接数値によつての表現ではなく量や色を用いた表現に対する人の認知特性の違い、および色に関するマッピングが判断に与える影響について調査することである。

本実験では、本務に相当する作業としてタイピングゲームを用いる。また、棒グラフにマッピングする人工的なデータの増加パターンとして、以下の 4 種類を用いる。データ数の上限は 10000 とする。

P1: 一定の速度で上昇変化する。

P2: 一定の速度で上昇変化する。P1 よりも高速。

P3: ランダムな速度で上昇変化する。

P4: ランダムな速度で上昇変化する。P3 よりも高速。

4.2. 実験 (1) 概要

本実験は、20 代の工学系大学生、大学院生 12 人を対象に行った。

実験の概要は以下のとおりである。

- I. タイピングゲーム『寿司打⁵』を難易度『高級コース: 練習』で行ってもらい、結果を記録する。
- II. 図 1 に示す棒グラフ形式のメータを起動する。タイピングゲーム『寿司打』難易度『お手軽コース: 練習』を行いながら、棒グラフの挙動、リセット動作の確認を一度だけ行ってもらう。

⁵<http://typing.sakura.ne.jp/sushida/>

III. タイピングゲーム『寿司打』難易度『高級コース：練習』を行いながら、棒グラフ形式のメータのリセットを行ってもらい、その結果を記録する。

IV. 実験中、リセットの際に意識していたことについてのアンケートを実施する。

タイピングゲームの実行時間は一回あたり約2分である。実験に用いたデータ増加パターン (P1~P4)、色パターン (図2の①~⑤) の組み合わせを表1に示す。実験協力者がリセット動作を行うたびに、実行順序に示す順番で組み合わせが変化する。

表1. データ増加パターン、色パターンの組み合わせと実行順序

		実験協力者 (A~L)			
		A,E,I	B,F,J	C,G,K	D,H,L
実行順序	1	P1, ①	P2, ②	P3, ③	P4, ④
	2	P3, ②	P4, ①	P1, ④	P2, ③
	3	P4, ③	P3, ④	P2, ①	P1, ②
	4	P2, ④	P1, ③	P4, ②	P3, ①

4.3. 実験 (1) 結果

表2は、各色パターンでのデータの蓄積量を実験協力者毎に示している。4.1節で述べたとおり、データ数の上限は10000であるため、データ蓄積量の最大値も10000となる。表2において10000と示されているものは、リセットをせずに最大値を超えてしまったものである。

表2. 実験 (1) 結果

		色パターン			
		①	②	③	④
実験協力者	A	9150	9170	9450	8908
	B	8884	8550	8901	9000
	C	9300	9658	9650	9314
	D	9917	9600	9718	9750
	E	9100	10000	8700	8796
	F	8281	7950	8741	8600
	G	9600	9755	9150	8880
	H	9313	9500	9421	8850
	I	8800	7944	8100	7914
	J	10000	9450	9201	8450
	K	9150	7976	8900	9610
	L	8941	9400	8378	8700

データ蓄積量を実験協力者、色パターン毎に箱ひげ図で表した結果を図3、図4に示す。

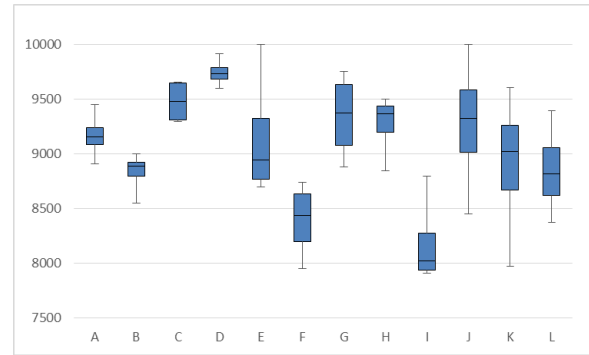


図3. 実験協力者毎のデータ蓄積量の分布

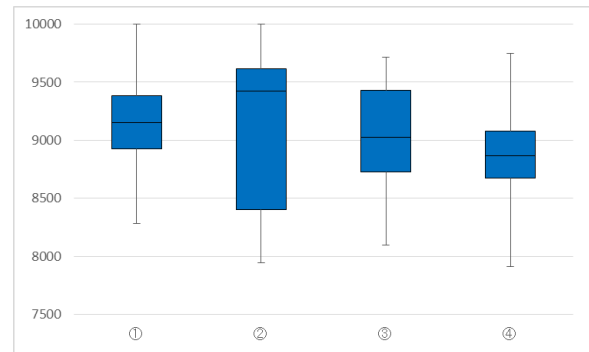


図4. 色パターン毎のデータ蓄積量の分布

実験協力者間のデータ蓄積量の比較では、優位水準5%で有意な差が見られたが、色パターン間では有意な差は確認されなかった。

アンケートの結果、棒グラフの色については全く意識していない実験協力者が多数であった。色パターンをデータ量の変化の確認に用いている実験協力者も存在したが、色パターンだけを基準にリセットを行う実験協力者は存在しなかった。

以上の結果より、データ蓄積量に対する確認タイピングはユーザによって異なるを考える。

4.4. 実験 (2) 概要

実験 (1) では、色パターン毎に一回ずつ実験を行ってもらっていた。これに対し、実験 (2) では、実験協力者毎に同じ色パターンで複数回実験を行ってもらい、リセット時データ蓄積量の平均やばらつきについて検証する。20代の工学系大学生、大学院生8人を対象に行った。

実験の概要は以下のとおりである。

- I. タイピングゲーム『寿司打』を難易度『高級コース：練習』で行ってもらい、結果を記録する。
- II. 図1に示す棒グラフ形式のメータを起動する。タイピングゲーム『寿司打』難易度『お手軽コース：練習』を行いながら、棒グラフの挙動、

- III. リセット動作の確認を一度だけ行ってもらう。タイピングゲーム『寿司打』難易度『高級コース：練習』を行いながら、棒グラフ形式のメータのリセットを行ってもらい、その結果を記録する。この手順は色パターン①で行う。
- IV. 色パターンを④あるいは⑤に変更し、手順IIIと同様の実験を行ってもらう。
- V. 実験中リセットの際に意識していたこと、色に対して意識していたことについてのアンケートを実施する。

IIIおよびIVに関して、実験に用いたデータ増加パターン、色パターンの組み合わせを表3に示す。実験協力者がリセット動作を行うたびに、実行順序に示す順番でデータ増加パターンが変化する。

表3. データ増加パターン、色パターンの組み合わせと実行順序

		実験協力者 (A~H)		
		全員 (III)	A,B,E,F (IV)	C,D,G,H (IV)
実行順序	1	P1, ①	P1, ④	P1, ⑤
	2	P3, ①	P3, ④	P3, ⑤
	3	P2, ①	P2, ④	P2, ⑤
	4	P4, ①	P4, ④	P4, ⑤

4.5. 実験 (2) 結果

一回のタイピングゲームの間に、すべての実験協力者は4回のリセットを行った。実験協力者 A,B,E,Fの色パターン毎のリセット時のデータ蓄積量の平均、標準偏差を表4に示す。また、実験協力者 C,D,G,Hの実験結果を同様に表5に示す。

表4. 実験 (2) 実験協力者 A,B,E,Fの各色パターンでの平均、標準偏差

		色パターン		
		①	④	
実験協力者	A	平均	9459.25	9756.25
		標準偏差	147.14	197.09
	B	平均	9922.50	9875.00
		標準偏差	106.18	107.61
	E	平均	9645.75	9397.25
		標準偏差	264.31	208.30
	F	平均	9422.25	9644.75
		標準偏差	424.99	62.57

表5. 実験 (2) 実験協力者 C,D,G,Hの各色パターンでの平均、標準偏差

		色パターン		
		①	⑤	
実験協力者	C	平均	9448.00	9407.75
		標準偏差	209.12	212.75
	D	平均	9776.25	9816.00
		標準偏差	105.86	36.18
	G	平均	9908.75	9835.75
		標準偏差	106.96	200.75
	H	平均	8201.50	8939.00
		標準偏差	412.51	242.34

実験協力者間でのリセット時のデータの蓄積量に関して、どの色パターンでも優位水準5%で統計的に有意な差が見られた。色パターン間の比較では、実験協力者毎の平均、および標準偏差いずれについても有意な差は確認されなかった。

アンケートの結果、④の色相環のように色が頻繁に変化すること、①の青色のように元々の彩度が高いものが動くことでタイピングゲーム中に頻繁に棒グラフに注意が向かってしまい、集中できなかったという実験協力者が存在した。

以上2つの実験結果より、ユーザ毎に棒グラフに対する確認タイミングの意識は異なり、ユーザ毎に可視化の調整を行う必要があると考える。

一方、色の変化は確認タイミングの変動の要因になりうる可能性があるが、ユーザによって反応や好みが異なり、過度な色の変化は本務に悪影響を与えてしまう可能性があると考えられる。

5. おわりに

本稿では、オンラインニュース等のストリームデータを確認するタイミングの判断を支援するシステムの構築に向けて、棒グラフ形式のメータを利用したユーザ実験を行った結果について報告し、棒グラフの色や個人差の影響について考察した。

ユーザ実験により、ユーザによって棒グラフの量に対する認知特性に有意な差がある結果が得られた。そのため、ユーザ毎に可視化の調整を行う必要があると考える。

一方、色の違いは確認タイミングの手がかりになるとは限らず、過度な色の変化、彩度の高い物体の変化は注意を集め、本務の集中を妨げる可能性があることがわかった。今後は、ユーザ毎の確認タイミングの差を調整するための手法の考案、可視化に用いる色の再検討を行い、インタフェースの開発を進める予定である。

参考文献

- [1] M. Czerwinski, E. Horvitz, and S. Wilhite: A diary study of task switching and interruptions, 2004 Conference on Human Factors in Computing Systems (CHI'04) , pp. 175-182 (2004)
- [2] B. O'Conaill, D. Frohlich: Timespace in the workplace: dealing with interruptions, 1995 Conference on Human Factors in Computing Systems (CHI'95), pp. 262-263 (1995)
- [3] 卓 璐, 王 琛, 浅井 洋樹, 山名 早人: 3 軸加速度を用いたデスクワーク中の割り込み可能性の推定, DEIM Forum 2015, E1-5 (2015)
- [4] 谷 堯尚, 山田 誠二: 机上にかかる圧力を用いたユーザの割り込み可能性推定, 人工知能学会論文誌, vol.29, No1, pp. 129-136 (2014)
- [5] 田中 貴紘, 藤田 欣也: オフィスワーカーの状況推定—割り込み拒否度を中心に—, 電子情報通信学会誌, Vol. 95, No. 5, pp. 457-460 (2012)
- [6] 沼野 航希, 高間 康史: オンラインニュースを対象としたモニタリングシステムの提案, 第 8 回インタラクティブ情報アクセスと可視化マイニング研究会, pp.18-23 (2014)
- [7] Y. Takama, M.Okumura: Interactive Visualization System for Monitoring Support Targeting Multiple BBS Threads, International Journal on Intelligent Decision Technologies, (DOI) 10.3233/IDT-140232 (2014)
- [8] Y.Takama, T.Kurosawa: Visualization System for Monitoring Bug Update Information, Trans.IEICE, Vol.E97-D, No.4, pp. 654-662 (2014)
- [9] R. Mazza (著), 加藤 諒 (編集), 中本 浩 (翻訳) , 情報を見える形にする技術, pp. 45-47 (2011)