

大規模ログデータの自動解析システム

三木 光範† 廣安 知之‡ 吉見 真聡† 松岡 拓也‡

†同志社大学理工学部 ‡同志社大学生命医科学部 ‡同志社大学工学部

1 はじめに

近年、センサネットワークや大規模な設備の稼働状況などに関する膨大なログデータが蓄積され続けている。これらのログデータを可視化することは、詳細な稼働状況を把握するうえで極めて重要である。そこで、本研究では大規模システムのログデータを可視化し、稼働状況の把握を行う手法について検討を行う。本研究で対象とした大規模システムは、著者らが開発を行っている知的照明システムである。

知的照明システムとは、オフィスワークにとって最適な光環境を提供することにより、知的生産性の向上と省エネルギーを目的としたシステムである [1]。本システムは膨大なログデータを解析することにより、管理者が知的照明システムの稼働状況を把握するために有益な情報を抽出し、グラフ化を行い、報告書を作成する自動解析システムである。作成したデータはウェブ上で管理することにより、ペーパーレス化の実現と開発者やユーザの情報共有を実現する。

2 知的照明システム

本研究で対象とする知的照明システムは共同研究先企業に導入を行った実証実験用の知的照明システムを対象とした。この知的照明システムは電球色蛍光灯 1 管、白色蛍光灯 2 管の合計 3 管を 1 組とする複数の照明器具と複数の移動可能な照度センサ、および電力計をネットワークに接続することによって構成される。オフィスワークが机上のパソコンから目標の明るさ (目標照度) を設定することにより、コンピュータが各照度センサの目標照度を満たす最適な点灯パターンに照明を制御する。3 管を 1 組の照明としているのは光の色 (色温度) を調節するためである。この照明システムはオフィスワークに対して最適な明るさと色温度を提供することにより、知的生産性の向上を図ることを目的とする。この知的照明システムでは、各オフィスワークの目標照度を満たすため、各オフィスワークの机上に照度センサを設置している。

知的照明システムは、照度センサが取得する現在照

Automated data analysis system
for large-scaled log data

† Mitsunori MIKI (mmiki@mail.doshisha.ac.jp)

† Tomoyuki HIROYASU

† Masato YOSHIMI

† Takuya MATSUOKA

Doshisha University (†‡)

度を基に各オフィスワークの目標照度を満たすよう照明が最適な明るさ (光度) を決定する [1]。

知的照明システムは Table.1 に示すログデータをシステムの稼働状況として毎分出力する。

Table. 1: ログデータの内容

項目	内容
出力日時	ログデータの出力日時 (年日時分)
白色光度	白色の照明が点灯する明るさ (単位は cd) × 照明台数
電球色光度	電球色の照明が点灯する明るさ (単位は cd) × 照明台数
現在照度	照度センサが取得した各机上面の明るさ (単位は lx) × 照度センサ数
目標照度	オフィスワークが設定した目標の明るさ (単位は lx) × 照度センサ数

3 ログデータの自動解析システム

解析対象とした知的照明システムは照明 26 台、照度センサ 24 台の構成となっており 1ヶ月間に出力されるログデータは 400 万以上になる。そのためこれらのログデータを手作業で解析することは不可能である。そこで、これらのログデータを解析するにあたり、管理者が解析を行うことを可能にすることを目的とし、管理者にとって有益であると考えられる情報を自動的に抽出する解析システムの構築を行った。

そこで膨大なログデータから管理者にとって有益であると考えられる情報を提供するため特徴抽出を行う。具体的には、知的照明システムの異常値や特異なデータが出やすいと予測される平均値の最大値・最小値および利用頻度が表れやすいと予測される変化量の最大値・最小値を特徴抽出データとして抽出する。これらのデータを抽出することにより、知的照明システムの導入効果やシステムトラブル等の検証、照度センサ等のハードウェアトラブル等を検出することが容易になると予測される。

さらに、可視化を行うため特徴抽出データを Gnuplot によってグラフ化することで知的照明システムの稼働状況の把握を容易にした。可視化したデータの例として、1日ごとの平均値をプロットした 1ヶ月間の消費電力量のグラフを Fig.1 に示す。このようにして生成したグラフは毎月定期的に作成する。特徴抽出データのグラフはデータ項目ごとに 1ヶ月間の日数分出力されるため、毎月 150 枚近くの特徴抽出データのグラフが生成される。したがって、これらの特徴抽出データのグラフを管理することは容易ではない。そこで、特徴抽出データのグラフを Tex に自動的に組み込み、これらのグラフを一冊の報告書として生成するプログラムを構築した。このプログラムにより、毎月の特徴抽

出データのグラフを管理することを容易にした。しかし、1ヶ月間の特徴抽出グラフは膨大であるため、すべての特徴抽出グラフを報告書に載せると管理が容易でないうえ、目的のグラフを探すことも容易ではない。

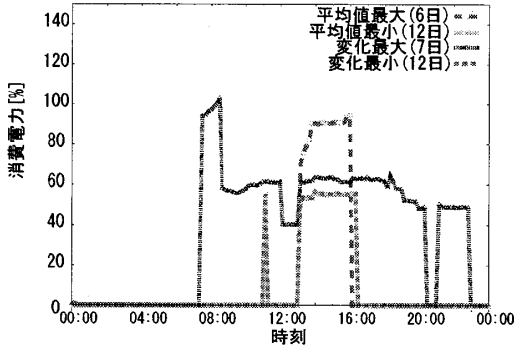


Fig. 1: 特徴抽出データを可視化した例

そこで紙の報告書には重要な情報のみを載せ、ユーザや開発者がログデータを閲覧するため、ウェブ上でログデータを管理するシステムを構築した。実際に構築したログデータ管理システムの UI 画面を Fig.2 に示す。ここで指定した条件の特徴抽出データのグラフが一括検索可能となっている。

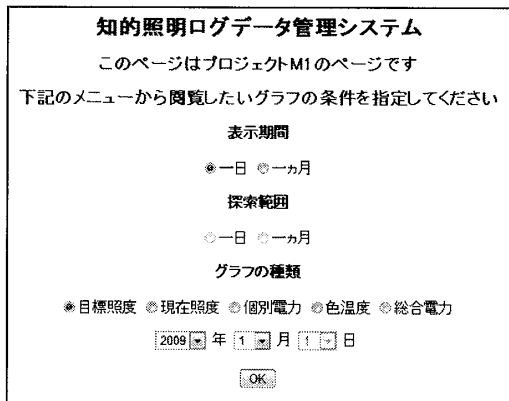


Fig. 2: ログデータ管理システムの UI 画面

このシステムを利用することにより、ウェブ上で特徴抽出データを一括に検索することが可能となりペーパレス化を実現するとともに、開発者やユーザの情報共有を実現する。さらに、特徴抽出データには表れないが、解析を行う上で知的照明システムのより詳細な稼働状況を把握する必要もあると考えられるため、照明もしくは照度センサごとにグラフ化を可能にする機能も実装した。この機能を利用することにより、特徴抽出によって得られた情報の多面的な検証が可能となる。

4 考察

構築した大規模ログデータの自動解析システムの効果の検証を行う。まず、Fig.3 に 1ヶ月間の現在照度の

特徴抽出によって生成されたグラフを示す。このグラフは 1 日ごとの現在照度の平均値を 1 カ月分プロットしている。

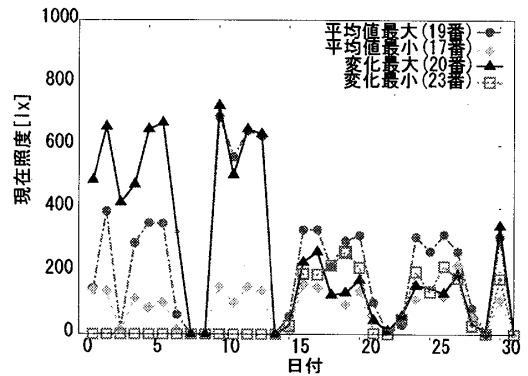


Fig. 3: 特徴抽出データ (目標照度)

Fig.3 の結果から変化量が最大であった人は目標照度を日によって大きく変えていることがわかった。

また変化量が最小であった人は平均値が 100 lx 程度であることから照度センサの設置場所が適切でなかった等の故障が起こっている可能性が考えられる。またウェブ上で詳細な稼働状況を確認したところ、オフィスの好む照度は 600 lx や 350 lx 等であることがわかった。この結果は一般的なオフィス環境で推奨されている照度である 750 lx[2] と比較して、実際のオフィスの好む照度が非常に低いことを示している。このことから、知的照明システムの個別照度の実現が、オフィスの光環境を改善することに効果的であることが確認できた。

これらのことから、本システムが知的照明システムの大量のログデータからユーザにとって有益な情報を提供していることが確認できる。

5 まとめ

近年、センサネットワークや大規模な設備などの稼働状況に関する膨大なログデータが蓄積され続けている。このため、ここで提案した方法を用いることでシステムの改良や故障発見が容易となる。

参考文献

- [1] 知的照明システムと知的オフィス環境コンソーシアム, 人工知能学会, Vol.22, No.3, pp.399-410, 2007
- [2] オーム社, 照明ハンドブック, 照明学会, 第 2 版, pp.270-271, 2003