様式 C-19

科学研究費補助金研究成果報告書

平成２２年　５月２８日現在

研究種目：基盤研究(B)
研究期間：2007～2009
課題番号：１９３００２８３
研究課題名（和文）学校ネットワークにおけるトラフィック情報マイニングを用いた異常検知システムの構築
研究課題名（英文）Designing of traffic mining system that detects an abnormal behavior in the school networking

研究代表者
金西　計英（KANENISHI　KAZUHIDE）
徳島大学・大学開放実践センター・教授
研究者番号：８０２０４５７７

研究成果の概要（和文）：Peer-to-Peer（P2P）型ファイル共有ソフトのもたらす問題は，もはや座視できない．多くの学校では，P2P ソフトの利用を禁止しているが，その利用を制限するのは困難である．そのため，ネットワークの管理者の負担は増大している．本研究では，ネットワークの異常検出に対し，管理者の支援を目指す．具体的には，トラフィックの可視化をおこなうトラフィックマイニングツールを開発し，その有効性を検証した．

研究成果の概要（英文）：Any longer, we cannot overlook the problem that the Peer-to-Peer(P2P) file-sharing software causes. The school staff prohibits the use of the P2P software at a lot of schools. However, it is difficult to limit the P2P software use. The load of the network administrator increases. In this research, we aim at the administrator's support for the anomaly detection on the network. We developed the traffic-mining tool that can visualize the traffic status concretely, and verified the effectiveness.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
２００７年度	3,800,000	1,140,000	4,940,000
２００８年度	3,500,000	1,050,000	4,550,000
２００９年度	3,000,000	900,000	3,900,000
年度			
年度			
総　計	10,300,000	3,090,000	13,390,000

研究分野：教育工学
科研費の分科・細目：科学教育・教育工学、　教育工学
キーワード：管理者支援、マイニング、SOM、可視化、異常検知、P2P 通信、ベクトル空間モデル

１．研究開始当初の背景

1990 年後半以降，インターネットが社会のインフラとなるにつれ，学校現場の情報化は大きく進んだ．学校現場にとって，インターネットは，授業だけではなく，必要不可欠なものとなっている．

情報インフラが適切に機能するために，情報インフラを運用する管理者の存在を抜きにすることはできない．管理者はネットワークの状態を監視し，異常が発生した場合，速やかに対処しなければならないが，ネットワークの利用が増えれば，それに比例して作業

量も増える．多くの場合，学校現場では，専門の管理者がいるわけではない．管理者は存在するものの，兼任であったり，名前だけであったり，必ずしも専門的な知識や経験を有しているわけではない．多様な学校ネットワークの管理者に対する支援は，切実に求められている．

管理者のタスクの主なものの一つに，ネットワークの監視がある．ネットワークのトラフィック上の異常という点で，Peer-to-Peer（P2P）通信に注目が集まっている．P2P 通信の引き起こす事件が，社会問題となっているからである．

P2P ソフトウェアの利用が問題になるのは，一つにネットワークの帯域の圧迫がある．P2P の利用によって，対外接続のネットワークの帯域が占有されてしまう．そのため，DNS，電子メール，HTTP 等の通信が阻害され，学校の様々な活動に影響が出る．

また，ファイル共有ソフトを用いて，音楽，映画，画像等，あらゆる種類のコンテンツが共有され，これらの中には著作権の保護対象である著作物を，違法に流通させているものが多い．著作権者の許諾を得ず，コンテンツを複製・配信することは大きな問題である．

しかし，より深刻な問題は，ファイル共有ソフトウェアの使用による情報漏洩のリスクである．Antinny などの暴露型コンピュータウイルスは，感染したコンピュータ内の様々なファイルを，ファイル共有ネットワークへ公開する．この結果，使用している PC のローカルに保存された情報が，ファイル共有ネットワークへ流出する．個人のプライベートな動画像，自衛隊や警察の機密資料等がファイル共有ネットワークに流出した情報漏洩事件が社会問題となったことは記憶に新しい．無論，ファイル共有ネットワークへの情報漏洩は，ファイル共有ソフトが直接の原因ではない．しかし，ファイル共有ソフトの利用が，暴露ウイルスの罹患を招き，情報漏洩のリスクを高めることは間違いない．

このように，学校現場でのインターネットの利用が拡大する中で，リスクが高まっている現状がある．管理者には，ネットワークの技術についての適切な知識，豊富な経験が求められる．とくに，経験に関する知識として，ネットワークの利用者が，ネットワークに利用においてどのように振る舞うかの深い理解が求められる．多くは，暗黙の経験則として明らかになることはない．ネットワークの管理者に対する支援として，タスクを軽減するツールの開発や，管理者の経験則の蓄積が求められている．

２．研究の目的

本研究の目的は，学校現場で何らかの形でネットワークの管理に携わっている管理者に対し，ネットワーク運用の負担を軽減するための手法について提案するものである．そのため本研究では，ネットワーク管理者のネットワーク監視作業を補完するツールの提供を目指している．具体的には，ネットワークの異常な振る舞いに関する調査と，通常とは異なるトラフィック上の振る舞いをシステムが発見するツールの開発をおこなう．本研究では，トラフィックからの異常発見手法を，トラフィックマイニングと呼び，トラフィックマイニングを実現するツールの提供を目指す．

ネットワークの様々な異常に関する調査の中から，P2P 通信が大きな問題となっていることが浮かび上がってきた．管理者が P2P ファイル共有ソフトを制限する場合，まず，パケットフィルタの利用が考えられる．しかし，P2P ノードは，不特定多数かつ可変であるため，IP アドレスによるフィルタリングは困難である．さらに，一部の P2P ソフトは，標準的な待機ポートを持たない．そのため，ポート番号によるフィルタリングも困難である．ファイル共有ソフトは，常に，管理ツール等の裏を掻く方法を求めており，管理者とファイル共有ソフト開発者とのいたちごっこの様相を呈している．

以上のような点を踏まえ，我々のトラフィックマイニングツールは，先ずは，学校現場で用いられているネットワーク内からの P2P 通信を検出することを目的とすることとした．無論，P2P に使用を限定するものではないが，大きな問題となっている P2P に焦点を当てることにした．

３．研究の方法

(1) 異常検知の支援

①トラフィックマイニングによる異常検知の支援

本研究では，トラフィックの異常発見を支援する枠組みの構築を目指している．通常，ネットワーク管理者は，自らが管理するネットワークの状態を，監視している．管理者によるネットワーク監視タスクは，主に，各種のログを収集し，収集したログを解析することが主な作業となっている．ログの多くは，ある種のフォーマットに従ったテキストである．つまり，管理者は，大量のテキスト情報の中から，何らかの意味を読み取るということをおこなっている．そこで，我々は，このログ情報をクラスタリングし，その上で可視化するようなツールを提供する．各種のログ情報を可視化することで，管理者にとって情報を解釈する負担が軽減される，と考えるからである．

我々の提案する手法は，トラフィックの全体の傾向を対象としている．ここでの主な処理対象は，TCP パケットのヘッダである．こ

れは，通信における個人のプライバシーに配慮した結果である．当初，我々は様々な情報を収集し，クラスタリングすることを目指していた．無論，多様な情報を集めることで，解析結果の精度の向上が期待された．一方で，多様な情報を集めることは，プライバシーの侵害に抵触する可能性も秘めている．我々は，最小の情報で，解析精度を高める方法を模索することにした．なお，多様な情報からの異常発見については，個人情報保護についての配慮しながら研究を続ける予定である．

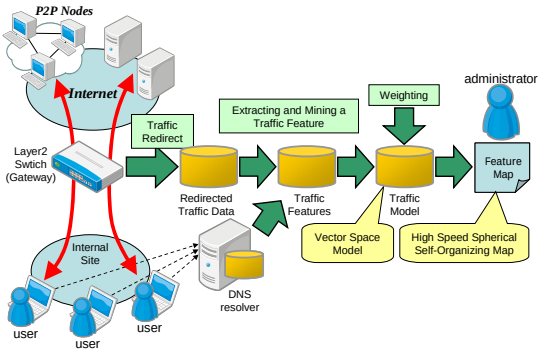


図 1. トラフィックマイニングの手順

図 1 に，異常検出の支援の枠組みについて示す．我々は，これら一連の流れから実現されるトラフィック特徴強化と可視化の流れ基づく異常発見支援の手法を，トラフィックマイニングと呼ぶ．その上で，トラフィックマイニングの支援ツールをトラフィックマイニングツールと呼ぶ．トラフィックマイニングツールでは，「トラフィックモデル生成」「特徴強化」「可視化」の処理をおこない，管理者の異常検知の発見を支援する．

②トラフィックモデルの生成

トラフィックマイニングをおこなうために，トラフィックマイニングツールでは，学内の送信トラフィックを収集し，トラフィックマイニングの基盤となるデータセットを作成する．このデータセットは，学校のネットワークにおけるトラフィックの全体傾向を表現したものである．あくまでも，全体を表現したものである点が重要である．

③特徴の強化

クラスタリング処理において，処理すべきデータに対しておこなう前処理は重要である．一般に，クラスタリングの前処理として，データの正規化と，特徴の強調が挙げられる．我々も，トラフィックモデルに対し，特徴量の強化をおこなう．

④可視化

可視化の過程では，集積されたトラフィックモデルから単位時間ごとに結果を可視化し，管理者に提示する．単に得られた情報を一括して可視化するのではなく，単位時間毎に可視化することで，トラフィック全体の俯瞰が可能となり，さらには変化の追跡も容易となる．

また，我々が情報の可視化を用いるのは，全体の傾向からの異常発見に重点をおくからである．トラフィックマイニングの手法では，個人のデータを扱うのではなくトラフィック全体の俯瞰図を提供する．

(2) 可視化の方法

①トラフィックモデルの構成

トラフィックモデルは単位時間におけるトラフィック特性を定量的に集積する．本研究では，ベクトル空間モデル（Vector Space Model:VSM）を採用した．トラフィックモデルは，一定時間のトラフィックの状態を表した特徴ベクトルである．特徴ベクトルはネットワーク中の個々の端末のトラフィック情報を表した要素ベクトルの集合として表される．ベクトルモデルは，ネットワークのトラフィック情報から，何らかの特徴を計算によって導き出す基盤を提供する．

②自己組織化マップを用いた可視化

生成されたトラフィックモデルは多次元ベクトル集合となっている．これは送信元 IP アドレスと宛先 IP アドレスの関係が，多次元空間上の分布として表現できることを意味する．人間は基本的に三次元までの空間は直感的に把握可能だが，それ以上の多次元空間を直感的に把握するのは困難である．

我々は，トラフィック情報を可視化するために自己組織化マップ(Self-Organizing Map:SOM)を用いる．SOM は，Kohonen らによって提案された 2 層のニューラルネットワークで構成される教師なし競合学習モデルである．本研究では，SOM を改良した球面 SOM を用いる．球面 SOM は，従来の SOM にあったエッジに配置されたノードと，その他のノードとの間の計算の不均衡の解消を目指した手法である．

SOM は，データ間の幾何学的構造を，可能な限り保ったまま二次元平面に写像する．同時にクラスタリングをおこなう．トラフィック状態を示した特徴ベクトルに SOM を適用することで，二次元に展開される．これを特徴マップと呼ぶ．特徴マップを提供することで，管理者は，自らが管理している組織のトラフィックの傾向を二次元平面上で俯瞰することが可能となる．

(3) トラフィックマイニングシステム

ここでは，システムの構成について述べる．システムの概要を図 2 に示す．

①トラフィック収集部

トラフィック収集部では，監視対象ネットワークのトラフィックの情報を収集・蓄積する．学校内のネットワーク機器と連動し情報を収集することを想定している．

②トラフィック解析部

トラフィック解析部では，収集したトラフ

ィック情報から特徴ベクトルを作成するための前処理をおこなう．収集された IP パケット群のヘッダから必要な情報を抽出する．

③DNS クエリ解析部

DNS クエリ解析部では，特徴ベクトル生成のため，DNS 関連の重み付けのための前処理をおこなう．

④モデル化部

モデル化部では，トラフィックモデルである特徴ベクトルを生成する．また，モデル化部では，この特徴ベクトルに対する重み付けを行う．P2P の特徴が表出するよう特徴ベクトルに対し，重み付けの処理をおこなう．P2P ソフトウェアは，一般に，接続する相手ノードの IP アドレスを直接指定する．このため DNS による名前解決が発生しない．この点に着目して重み付けをおこなうことにした．

⑤可視化部

得られた特徴ベクトルへ，SOM アルゴリズムを適用することで特徴マップが生成される．ここで生成された特徴マップにはトラフィックの全体状態が表現される．管理者は特徴マップを眺め，ネットワーク上の何らかの特異トラフィックの存在に気付くことができる．

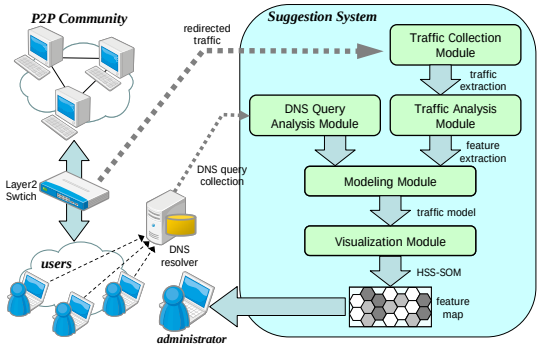


図 2. トラフィックマイニングシステムの概要

４．研究成果

(1) 可視化の検証実験と考察

我々はトラフィックマイニングツールを試作し，提案手法の評価をおこなった．評価は，システムによる可視化の性能を調べるというものである．

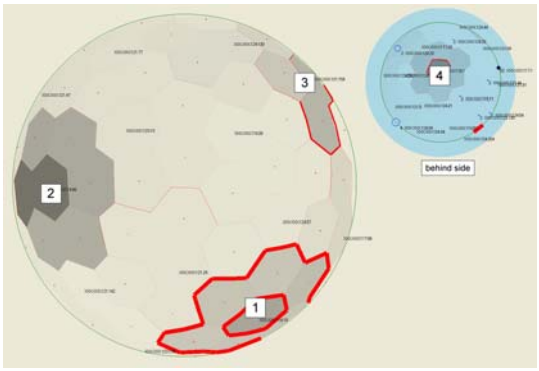
以下に，成果の一つとして 2008 年 9 月の実験の結果を示す．この実験ではある組織の協力を得て，当該組織のネットワークの 1 時間分のトラフィック情報を収集し，これを実験データとした．実験の基となったデータは，1 時間のパケット総数が 2,329,730 件であり，特徴ベクトル中の要素ベクトルは 26,480 件であった．

実験中 1 台の端末において意図的に P2P ファイル交換ソフト(Share)を動作させ，任意のデータファイルをダウンロードした．生成された特徴マップにおいて P2P のトラフィックが，管理者が注目できるような形で表出されているかどうかが問題となる．図 3 に今回の実験で生成した特徴マップを示す．

この特徴マップでは，大規模な通信をおこなっている端末がクラスタとして表出している．意図的に発生させた Share のトラフィックが，クラスタとして表出していることが分かる（赤く囲った部分，なお，赤い線は後から引いたもの）．Share に関係するトラフィックを，明確なクラスタとして表出させることができることが確かめられた．

実験にあたって，大学等のネットワークの管理に携わっている管理者にマップを評価してもらったところ，いずれも P2P トラフィックの存在について理解することができた．このことは，特徴マップに一定の有効性があることを示していると考える．システムが提示するマップの解釈は主観的な問題であり個人差が存在するものの，結論として，本研究で提案するトラフィックマイニングツールには，一定の可能性があるものといえる．

図 3. 特徴マップの例



(2) 研究の位置付けと展望

本研究は，ネットワークの管理者が，ネットワークの異常を検知するトラフィックマイニング手法を提案し，トラフィックマイニングツールの有効性を検証するものである．

そのため，我々は，まずネットワーク上の利用者の振る舞いに関して，様々な調査をおこなった．SNS 等，ネットワークでは新しいツールが生まれ，新しいコミュニケーションのチャンネルによって利用者の振る舞いの特徴が現れるからである．さまざまな通信の中からネットワークの異常検知という点で，P2P トラフィックに着目した．

P2P トラフィックは，既存のフィルタリング技術で制限することが困難である．何らかの新しい手法の開発が急がれる．我々は，トラフィックマイニングを提案する．提案の有効性を検証するために試作システムを構築し，特徴マップを作成した．トラフィックマイニングは，我々の実験結果から，P2P 通信を発見する上で，有効な手法となると考える．

今回の実験において，P2P 通信の中で，ファイル共有ソフトウェアのトラフィックのみを分離し可視化することはできなかったものの，P2P トラフィックと非 P2P トラフ

ィックとを比較的明確に分離表示することができた．P2P 通信については，クラスタリングが可能であることが分かった．

ネットワーク管理者の負担は，今後も増加すると予測される．管理者の作業を支援するようなツールの整備は，要求が高いと考える．今回の開発した手法は，結果的に，P2P に特化したものとなっているが，今後は，他のさまざまな問題の兆候についても提示できるように，モデルを洗練させる必要がある．そのためには，システムの改良と，ネットワーク上の利用者の動向の調査を，継続する必要がある．

５．主な発表論文等

〔雑誌論文〕（計 20 件）

① Satoshi Togawa 、 Kazuhide Kanenishi and Yoneo Yano 、 Peer-to-Peer File Sharing Communication Detection System with Traffic Mining and Visualization 、 Proceedings of 13th International Conference on Human-Computer Interaction (HCII2009)、査読有、LNCS 5610、2009、900-909

② 金西 計英、戸川 聡、松浦 健二、光原 弘幸、矢野米雄、ネットワークの利用動向からの異常検知手法について、学術情報処理研究、査読有、Vol.13、2009、74-83

③ 金西 計英、 松浦 健二、 大家 隆弘、 佐野 雅彦、 北 研二、 矢野 米雄、徳島大学におけるポータルシステムの構成とその運用について、大学情報システム環境研究、査読有、Vol.12、2009、34-42

④ Masahiro Nakagawa 、 Kazuhide Kanenishi 、 Kenji Matsuura 、 Yasuo Miyoshi、Hiroyuki Mitsuhara and Yoneo Yano、Authentication and Authorization exchange for University Federation、 Proceedings of the 17th International Conference on Computers in Education、査読有、-、2009、477-479

⑤ Hiroyuki Mitsuhara、Junko Matsumoto、Noriko Uosaki、Mihoko Teshigawara、Kenji Kume and Yoneo Yano 、 Niche-Learning: New Learning Style Using Public Display System 、 Proceedings of ED-MEDIA 2009 World Conference on Educational Multimedia, Hypermedia & Telecommunications、査読有、-、2009、1167-1175

⑥ Hiroyuki Mitsuhara、Yoshiki Yamada、Toshiyuki Moriyama 、 Kazuhide Kanenishi and Yoneo Yano、Paper-Top Interface in Classroom、Proceedings of International Conference on Cognition and Exploratory Learning in Digital Age 2009、査読有、-、2009、299-301

⑦ Yasuo Miyoshi、Ryo Okamoto、Kazuhide Kanenishi and Yoneo Yano、A Design of Social Networking Service for Supporting Learning Habits Development 、 Proceedings of International Conference on Cognition and Exploratory Learning in Digital Age 2009、査読有、-、2009、396-399

⑧ Shiho Fujii、Satoshi Togawa、Kazuhide Kanenishi and Yoneo Yano 、 Raising Hand Action Detection using Wireless Game Controller for Teaching Assistance System、2009 International Workshop on Nonlinear Circuits and Signal Processing (NCSP'09)、査読有、-、2009、325-328

⑨ 金西 計英、松浦 健二、三好 康夫、高木 知弘、嵯峨山 和美、矢野 米雄、大学間 WEB サービス連携のための Shibboleth を用いた許可管理機能の実現、日本教育工学会論文誌、査読有、Vol.32、2008、93-96

⑩ Kenji Matsuura、Naka Gotoda、Keiji Niki、Kazuhide Kanenishi and Yoneo Yano、Supporting multi-step annotation to promote reflective learning: triggered by a cell-phone、International Journal of Mobile Learning and Organization、査読有、Vol.2、2008、119-132

⑪ Kazuhide Kanenishi、Kenji Matsuura、Yasuo Miyoshi、Kazumi Sagayama 、 Tomohiro Takagi and Yoneo Yano 、 Design of Authentication Infrastructure for the WEB Service Federation between Universities 、 Proceedings of Association of Pacific Rim Universities 9th Distance Learning and the Internet Conference 2008、査読有、-、2008、255-262

⑫ Hiroyuki Mitsuhara 、 Kazuhide Kanenishi and Yoneo Yano、Handheld Review: Ubiquitous Technology-Based Method to Bridge Class and e-Learning、Proceedings of ICCE2008、査読有、-、2008、406-411

⑬ Kenji Matsuura、Kazuhide Kanenishi、Yasuo Miyoshi、Kazumi Sagayama and Yoneo Yano、Promoting Physical Skill Development in a Video-Based WEBlog Community Environment、Proceedings of ED-MEDIA2008、査読有、-、 2008 、936-945

⑭ 松浦 健二、仁木 啓司、後藤田 中、金西 計英、矢野 米雄、スライド教材の編集課題による教育・学習支援環境の試作、電子情報通信学会論文誌(D)、査読有、

Vol.J91-D、2008、259-268

⑮ Hiroyuki Mitsuhara、Kitamura Akihiro、Kazuhide Kanenishi and Yoneo Yano、Knowledge Trading Environment using Virtual Money: Approach to Motivating to Share Knowledge 、 International Journal of WSEAS Transactions on Information Science and Applications、査読有、Vol.2、2007、309-316

⑯ Hiroyuki Mitsuhara、Shunsuke Nakaya、Kazuhide Kanenishi and Yoneo Yano、E-Notebook Tool for Effective Knowledge Construction from Web: Encouragement of Multi-Perspective Thinking and Prevention of Copy-and Paste 、 Proceedings of the Seventh IASTED International Conference on Web-Based Education、査読有、-、2007、406-411

⑰ Hiroyuki Mitsuhara、Akihiro Kitamura、Masato Akatsuka 、 Naofumi Ise 、 Yasumoto Hirakawa 、 Kazuhide Kanenishi and Yoneo Yano、Knowledge Trading Environment and its Application to Student-Created e-Learning Material 、 ICCE2007 Supplementary Proceedings、査読有、-、2007、63-64

⑱ Yasuo Miyoshi、Katsuhito Nakagawa、Kazuhide Kanenishi 、 Kazumi Sagayama and Yoneo Yano、A Design and Prototyping of a Social Chronology System for History Learning、ICCE2007 Supplementary Proceedings、査読有、-、2007、303-308

⑲ Kazuhide Kanenishi、Kenji Matsuura、Hiroyuki Mitsuhara、Yasuo Miyoshi、Junko Minato and Yoneo Yano、Using the Peer Review between Students in the Collaborative Report Writing System 、 ICCE2007 Supplementary Proceedings、査読有、-、2007、23-24

⑳ Satoshi Togawa、Kazuhide Kanenishi and Yoneo Yano 、 Peer-to-Peer File Sharing Communication Detection System Using Network Traffic Mining、Proceedings of 12th International Conference HCI International 2007、査読有、-、2007、769-778

〔学会発表〕（計 9 件）

① 山崎　雄大、サラウンディングキャンパスにおける利用者コンテキストを活用した情報提供、教育システム情報学会平成 21 年度第 5 回研究会、平成 22 年 1 月 23 日、東北大学

② 金西　計英、大学間 Web サービス連携における間接的な認可の制御について、平成 21 年度情報教育研究集会、平成 21 年 11 月 9 日、東北大学

③ 金西　計英、トラフィックマイニングによるネットワークの異常検知について、日本教育工学会第 25 回全国大会、平成 21 年 9 月 21 日、東京大学

④ 金西　計英、学習コミュニティへの SNS による支援環境の構築について、教育システム情報学会第 34 回全国大会、平成 21 年 8 月 19 日、名古屋大学

⑤ 三好　康夫、習慣化支援 SNS の設計に向けた事前調査、教育システム情報学会第 34 回全国大会 、平成 21 年 8 月 19 日、名古屋大学

⑥ 中川　雄仁、ソーシャルブックマークとの連携による年表推薦機能を持つ歴史年表作成システムと評価、電子情報通信学会教育工学研究会、平成 21 年 3 月 7 日、香川大学

⑦ 中川　真宏、認証基盤を用いた Web システム間の連携の検討、平成 20 年度情報教育研究集会、平成 20 年 12 月 13 日、北九州国際会議場

⑧ 金西　計英、グループレポートの相互評価導入による作文能力の向上、日本教育工学会第 24 回全国大会、平成 20 年 10 月 11 日、上越教育大学

⑨ 松浦　健二、SNS 援用環境の研究を通じたコミュニティ型学習の一考察、日本教育工学会第 24 回全国大会、平成 20 年 9 月 3 日、熊本大学

６．研究組織

(1) 研究代表者

金西　計英（KANENISHI　KAZUHIDE）
徳島大学・大学開放実践センター・教授
研究者番号：８０２０４５７７

(2) 研究分担者

戸川　聡（TOGAWA　SATOSHI）
四国大学・経営情報学部・講師
研究者番号：２０３９９１６６

妻鳥　貴彦（MENDORI　TAKAHIKO）
高知工科大学・工学部・講師
研究者番号：６０３２０１２３

松浦　健二（MATSUURA　KENJI）
徳島大学・高度情報化基盤センター・准教授
研究者番号：１０３６３１３６

光原　弘幸（MITSUHARA　HIROYUKI）
徳島大学・大学院ソシオテクノサイエンス研究部・講師
研究者番号：９０３６３１３４

三好　康夫（MIYOSHI　YASUO）
高知大学・教育研究部自然科学系・助教
研究者番号：２０３８０１１５