

Windows アプリケーションのレスポンス時間 認識・予測ソフトウェアの開発 (ユーザ中心型性能予測ツール)

Response Time Recognition and Prediction Scheme for Windows Application

堀井 洋 1)
Hiroshi HORII

1) 早稲田大学理工学研究科情報科学専攻 (〒169 - 8555 東京都新宿区大久保 3 - 4 - 1 51 - 11
- 05 E-mail: horii@yama.info.waseda.ac.jp)

ABSTRACT. The performance of PC and purpose of using PC has been diversified significantly, and it gets more essential to acknowledge the most efficient architecture to optimize PC's performance according to the each usage. However using ordinary computer performance evaluation (CPE) based on Bench Mark method, it is impossible to evaluate PC performance level reflecting each usage.

We developed ResAnalyzer which is capable to acknowledge the response time of user action on Windows application and estimate the response time of same user action on each PC. Added to that it is also capable to determine priority level of each user action and derived from analyzing priority level and response time, it can calculate PC performance level. By using this ResAnalyzer, a user can choose a PC which is most efficient and has most optimized architecture.

In this paper, we tested on accuracy of acknowledgement of user action and its estimated response time of user action. As a result, it is capable to acknowledge 80% of major user action on Windows office and capable to estimate CPU time as a response time with less than 20% error.

1. はじめに

汎用 CPU の処理速度の向上, Memory のバンド幅の拡大・バイト単価の安価化等により, 市場で販売される PC は, 年々性能が向上している. また処理性能の高い PC の他にも, ノート PC や PDA など, 携帯性, 低電力性を重視した PC も販売されている.

一方, PC の性能向上・多様化の流れを受け, ユーザーの PC 利用用途も, オフィス系ソフトの利用, ビデオなどの AV 編集, Web・メールといったインターネットの利用など, 多様化してきている.

PC の性能向上・多様化, ユーザーの利用用途の多様化により, ユーザーは, 自身の PC 利用用途に最適な処理能力を持つ PC を選択する必要がある. しかし, 利用用途に最適な PC を選択するには, PC の性能を評価する手法が必要である.

PC の性能を評価する手法として, 現在, ベンチマークソフトを利用した手法が多く利用されている. 科学計算系ベンチマークソフトとして, SPEC[1], NAS Parallel Benchmarks[2], Himeno Bench[3], また, アプリケーションベンチマークとして, Winstone [4], SysMark[5]などが, PC の性能評価に, 多く利用されている. しかし, これらのベンチマークソフトを用いた性能評価手法は, 実際のユーザーが感じる PC の体感速度に直結した評価値を示すことはできない.

本稿では, ユーザーのアプリケーションに対する要求(ユーザーアクション)を, アプリケーションが処理終了するまでの時間(レスポンス時間)を比較することで, PC の性能を評価する手法を提案する. 本手法を用いることで, ユーザーにとって直感的で, わかりやすい PC の性能評価を行うことが可能である.

我々は, 本提案手法を実現する, ResAnalyzer の開発を

行った. ResAnalyzer は, (1)ユーザーアクションのレスポンス時間の認識, (2)ユーザーアクションの解析, (3)ユーザーアクションのレスポンス時間の予測, を行う.

これより, 第 2 節にてベンチマークソフトを用いた性能評価手法の問題点について, 第 3 節にて提案する PC 性能評価手法について, 第 4 節にて ResAnalyzer について, 第 5 節にて, ResAnalyzer の検証を行い, 第 6 節にてまとめを行う.

2. 従来手法の問題点

PC の性能を評価する手法として, ベンチマークソフトを利用する手法が多く利用されている.

ベンチマークソフトとは, ハードウェア, ソフトウェアの性能を評価するためのアプリケーションであり, ベンチマークソフトを利用した PC の性能評価は, 一般的に, ベンチマークプログラム内で行われる, 特定の処理を性能評価の対象となる PC で計算し, その実行時間を比較することで行われる.

ベンチマークソフトは, (1)デバイスベンチマーク, (2)アプリケーションベンチマークに分類することができる. デバイスベンチマークは, PC を構成する特定のデバイスの性能を評価するためのベンチマークソフトである. また, アプリケーションベンチマークは, 実際に利用されているアプリケーションを用いて PC を評価するベンチマークソフトである.

これより, デバイスベンチマーク, アプリケーションベンチマークについて, その特徴, 問題点を述べる.

2.1. デバイスベンチマーク

デバイスベンチマークが対象とするデバイスは, CPU, メモリ, グラフィックカード, サウンドカード, ハードディスク, ネットワーク, などが挙げられる. CPU, メモリを対象とするベンチマークソフトは, 特定のアルゴリズムを計算させた際の, CPU の MIPS, FLOPS を, グラフィックカードを対象とする

ベンチマークソフトは、2次元、3次元画像の描画速度を、サウンドカードを対象とするベンチマークソフトは、デバイスドライバの性能、ノイズの大きさなどを、ハードディスクを対象とするベンチマークソフトでは、ハードディスクの読み込み・書き込み速度を、ネットワーク・インタフェース・カードを対象とするベンチマークソフトでは、バンド幅、スループット、レイテンシーなどを測定する。Sisoftware Sandra [6]、3DMark[7]、WinBench[8]などがデバイスベンチマークとして広く利用されている。

デバイスベンチマークソフトは、(1)デバイスごとの性能を評価できる、(2)容易に評価可能、(3)比較基準が明確、といった利点がある。ユーザーは、デバイスベンチマークを利用することで、PC全体の性能を予測することが可能である。

しかし、デバイスベンチマークを用いた評価から、実際にユーザーが利用しているアプリケーションの処理速度を求めることはできない。それは、デバイスベンチマークの測定結果は、特定のデバイスのみには負荷がかかる場合が多く、一般的なユーザーが利用するアプリケーションにおける、様々なデバイスを組み合わせた処理性能を評価することはできないことが原因である。つまり、デバイスベンチマークでは、ユーザーに直感的なPCの性能値を与えることができない。

2.2. アプリケーションベンチマーク

アプリケーションベンチマークは、一般的なユーザーが利用するアプリケーションの処理時間を測定し、その処理時間からPCの性能評価を行う。Winstone、SysMarkなどがアプリケーションベンチマークとして広く利用されている。

WinStoneは、Microsoft社とZiff-Davis社が共同で開発したアプリケーションベンチマークで、実際に利用されているアプリケーションに一定の処理を行わせ、その処理時間から相対性能を導き出す。

SysMarkは、一般的なアプリケーション、OSの処理性能を評価するベンチマークを開発する非営利団体BAPCO(Business Applications Performance Corporation)によって開発されたアプリケーションベンチマークである。BAPCOには、Microsoft、IBM、AMD、Intel、Toshibaなどが加盟しており、現在の最新バージョンであるSysMark 2002に利用されるアプリケーションには、Microsoft Word 2002、Microsoft Excel 2002、Microsoft Power Point 2002などのOffice系アプリケーションの他に、Netscape Communicator、McAfee VirusScan、Winzipなどのアプリケーションも含まれている。

これらのアプリケーションベンチマークは、ユーザーが利用するアプリケーションを用いてPCの性能評価を行うことにより、ユーザーの体感的なPC性能を求めることができる。また、様々なアプリケーションベンダ、ハードウェアベンダが開発に携わっているため、他のデバイスベンチマークより信頼性の高い評価を行うことができる。

しかし、これらのアプリケーションベンチマークは、ユーザー固有のアプリケーションの利用方法に依存したPC性能評価値を求めることができない。実際にユーザーがPCを利用する際、ユーザーによって、アプリケーションの利用方法は異なる。例えば、常にバックグラウンドで音楽を流しながら、オフィス系アプリケーションを利用するユーザーや、常に2つ以上のアプリケーションを利用しているユーザーなどが挙げられる。アプリケーションベンチマークでも、複数のアプリケーションを同時実行させることで、PCの性能を評価することを行っているが、実際のユーザーの利用方法を網羅する

ことはできない。

また、アプリケーションベンチマークは、最新のアプリケーションや、測定対象となっていないアプリケーションを基にしたPCの性能評価はできない。例えば、グラフィック系アプリケーションなどを多用するユーザーにとっては、現状のアプリケーションベンチマークから適切なPC性能を得ることはできない。

このように、アプリケーションベンチマークでは、ユーザー固有のPC利用方法に対するPC性能評価を行うことができない。今後、多様化するPC性能、PC利用方法が生じる中、様々なユーザー層、様々なアプリケーションに対応するPC性能評価手法が求められている。

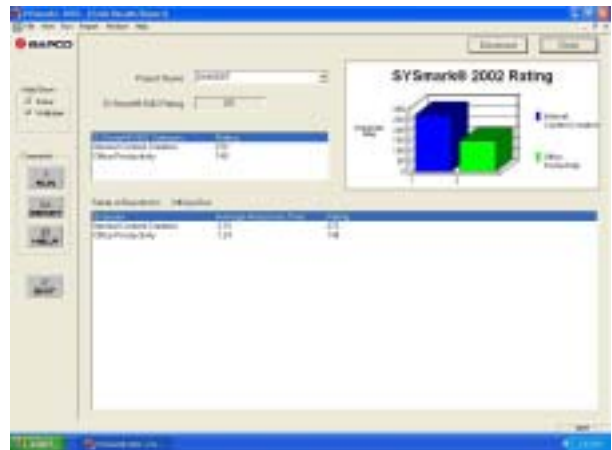


図 1: Sysmark2002

3. 提案する PC 性能評価手法

本稿で提案するPC性能評価手法は、ユーザーが利用しているPC環境で、ユーザーが利用しているアプリケーションのユーザーアクションのレスポンス時間を基に、PCを評価する。式1に、提案するPC性能評価手法を示す。

$$E = \sum_{i=1}^n (Trsp(i) \times w(i)) \quad \dots \text{式 1}$$

Eは、PC性能評価値を表し、Trsp(i)は、ユーザーアクションiのレスポンス時間を示す。また、w(i)は、ユーザーアクションiの重要度を示す。ユーザーアクションの重要度とは、ユーザーのユーザーアクションに対する重要度を示す。

本評価手法により計算されるPC性能評価値は、以下の特徴・問題点を持つ。

特徴

- ユーザー固有のPC利用方法に合わせたPC性能評価を行う。
- 全てのアプリケーションを対象に、PC性能評価を行う。

問題点

- ユーザーは、アプリケーションのレスポンス時間を測定することができない。
- 同じユーザーアクションを、複数のPCで実行するのは困難である。
- ユーザーアクションの重要度を求める必要がある。

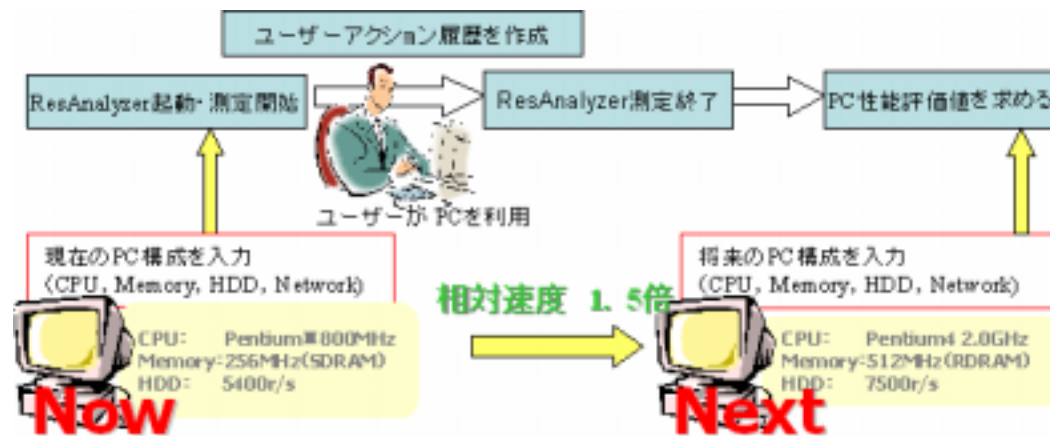


図 2: ResAnalyzer の利用方法

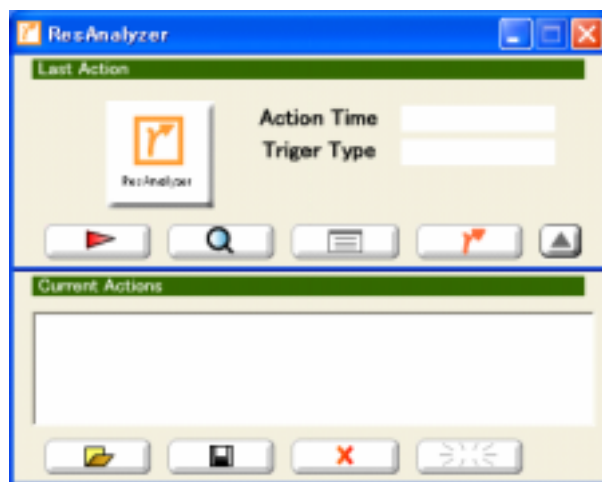


図 3: ResAnalyzer

4. ResAnalyzer

我々は、第 3 節にて提案した PC 性能評価手法を実現するアプリケーション、ResAnalyzer を開発した。これより、ResAnalyzer の概要を述べ、それぞれの機能について、記述する。

4.1. 概要

ResAnalyzer は、Windows より (1) ウィンドウメッセージデータ、(2) パフォーマンスデータ、を取得し、以下の機能を実現する。

- Windows 上で起動しているアプリケーションに対するユーザーアクションのレスポンス時間を認識
- 認識したユーザーアクションを、ユーザーが他の PC 上で実行した場合のレスポンス時間を予測
- アプリケーションの利用頻度を求めることで、ユーザーのユーザーアクションに対する重要度も求める

ResAnalyzer の利用方法を図 2 に示す。ResAnalyzer にて、PC の性能評価を行う場合、まず、ResAnalyzer を起動し、測定開始ボタンを押す。そして、ResAnalyzer が測定中に、ユーザーは、普段通りの PC 利用を行う。ユーザーが PC 利用を終了する際、ResAnalyzer の測定を停止し、PC の性能評価値を求める。なお、PC の性能評価値を求める際、将来の PC 構成の選択を行い、将来の PC 構成での

PC 性能評価値を求めることができる。ユーザーは、いくつかの将来の PC 構成での性能評価値を求めることにより、ユーザーにとって最適な PC 構成を選択することが可能となる。

4.2. ウィンドウメッセージデータの取得

ウィンドウメッセージは、Windows がウィンドウを所有するアプリケーションに対して送るメッセージのことである。マウスのクリック、キーボードのタイピングなどユーザーが行う全てのアクションは、ウィンドウメッセージを通じてアプリケーションに通知される。また、ボタンを押したり、スクロールバーを移動させたりするアクションは、Windows 側でそれらのイベントが処理され、アプリケーション側では、押されたボタンとその状態、スクロールバーの位置といったことがわかるようなメッセージが通知される。

ResAnalyzer では、Windows からアプリケーションに送られるウィンドウメッセージをフックする Hook.DLL を利用し、ウィンドウメッセージを取得する。

4.3. パフォーマンスデータの取得

Windows では、アプリケーション、サービス、ドライバーのチューニングのために、CPU、メモリ、ディスクなどの利用状況をパフォーマンスデータとして提供している。

ResAnalyzer では、Windows からパフォーマンスデータを取得する PerfReg.DLL を利用し、パフォーマンスデータを 0.1 秒¹ごとに取得する。PerfReg.DLL が取得するパフォーマンスデータより得られる値を、以下に示す。

Process

- CPU 使用率
- IO デバイスへの読み込み・書き込みサイズ
- ページフォルト数
- ワーキングセットサイズ
- 仮想メモリサイズ
- 起動時間

¹ パフォーマンスの取得を 0.1 秒以下の間隔で行うと、他アプリケーションへの挙動の影響が大きくなるため。



図 4: ユーザーアクションの PC 利用部分の遷移グラフ

Disk

- Disk 読み込み・書き込み使用率
- Disk 読み込み・書き込みサイズ

Memory

- 利用可能メモリサイズ
- ファイルキャッシュサイズ
- ファイルキャッシュミス数

Network Interface Card

- 送信・受信バイト数
- 現在のバンド幅

CPU 使用率や, Disk 読み込み・書き込み使用率は, 2 つのパフォーマンスデータを取得した時間中の CPU 使用率を示す。

4.4. ユーザーアクションのレスポンス時間認識

ResAnalyzer は, Windows 上で動く, 全てのアプリケーションに対するユーザーアクションのレスポンス時間を, 動的に認識する。ResAnalyzer は, 以下の開始条件, 終了条件を満たす区間を, レスポンス時間と特定する。

開始条件

キーボード, マウスボタンが動作後, 0.5 秒以内¹に WM_COMMAND, WM_NOTIFY コマンドを, レスポンス時間の特定対象となるアプリケーション プロセスが受信する。

終了条件

ユーザーアクションを実行しているアプリケーション プロセスの CPU 使用率が, ユーザーアクションを実行していない時の使用率と同じであり, かつ, Disk 及び Network を利用していない。

WM_COMMAND, WM_NOTIFY コマンドは, Windows からアプリケーションに送られるウィンドウメッセージであり, ユーザーがメニュー, ボタン, ショートカットキーのアクションを行う際に発行される。

4.5. ユーザーアクションのレスポンス時間の予測

ResAnalyzer は, 認識したユーザーアクションのレスポンス時間中のパフォーマンスデータを利用し, 他の PC で同じユーザーアクションを実行した際の, ユーザーアクションのレスポンス時間を予測する。

ユーザーアクションのレスポンス時間の予測は, (1) ユーザーアクションの解析, (2) CPU, Disk, Network ごとの利

用時間予測, (3) ユーザーアクション全体のレスポンス時間の予測, の行程で行われる。これより, 上記 3 つの行程を説明する。

4.5.1. ユーザーアクションの解析

アプリケーションは, ユーザーアクションを処理中, (1) CPU, (2) Disk, (3) Network, の使用する。ユーザーアクションを処理中, そのレスポンス時間は, これらの使用時間のいずれかに当てられている。ResAnalyzer は, ユーザーアクションの PC 使用部分の遷移を解析する。

ユーザーアクションの解析に当たり, ResAnalyzer は, ユーザーアクションのレスポンス時間内に取得した, パフォーマンスデータを利用する。利用するパフォーマンスデータを以下に示す。

- プロセスごとの CPU 使用率
- 全体の Disk 読み込み・書き込み使用率
- 全体のネットワーク送信・受信バイト数

これらのパフォーマンスデータから, 以下の時間情報を計算する。

- プロセスごとの CPU 使用時間
- プロセスごとの Disk 読み込み・書き込み時間
- プロセスごとのネットワーク送信・受信時間

プロセスごとの CPU 使用時間は, パフォーマンスデータを取得した時間と, パフォーマンスデータより求めることが可能である。プロセスごとの Disk 読み込み・書き込み時間も, プロセスごとの CPU 使用時間と同様にして, 全体の Disk 読み込み・書き込み時間を求め, プロセスごとの IO 書き込み・読み込みサイズより, プロセスごとの Disk 読み込み・書き込み時間を求める。プロセスごとのネットワーク送信・受信時間も Disk の読み込み・書き込み時間と同様に, プロセスごとの IO 書き込み・読み込みサイズより, プロセスごとのネットワーク送信・受信時間を求める。なお, 全体の送信・受信使用率は, 最大サイズの送信・受信を行っている場合を 100%とし, サイズに比例して使用率を求めるものとする。

ResAnalyzer は, プロセスごとの CPU 使用時間, プロセスごとの Disk 読み込み・書き込み時間, プロセスごとのネットワーク送信・受信時間を求めた後, これらの時間でレスポンス時間が埋まる様に配置する。また, CPU 使用時間, プロセスごとの Disk 読み込み・書き込み時間, プロセスごとのネットワーク送信・受信時間の開始時点と, その PC 利用部分の前に利用していた部分との間に, 依存関係を作成する。図 5に PC 利用部分の配置方法について示す。

¹ 実際に測定して, 最適な時間が 0.5 秒であった。

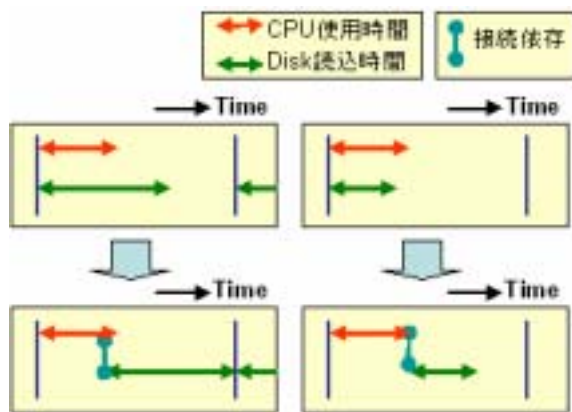


図 5: PC 利用部分の依存関係の解析

図 5は、CPU 使用時間と Disk 読み込み時間との間で、依存関係を作成する例である。それぞれの時間で、パフォーマンスデータを取得した2つの時刻の間を埋めるように配置する。この際、1つの利用部分が、最も長くなるように配置する。図 5の左の例の場合、次の時間帯に Disk の読み込みが行われているため、Disk の読み込み時間を右側にシフトする。

ResAnalyzer は、このように全てのレスポンス時間を埋めるように、PC 使用時間を配置し、図 4の様なグラフを作成する。

4.5.2. CPU, Disk, Network ごとの利用時間予測

ResAnalyzer は、4.5.1で求めた、それぞれの CPU, Disk, Network 使用時間ごとに、将来の PC 上での CPU, Disk, Network のそれぞれの使用時間を求める。

将来の PC 上での CPU の使用時間の予測は、それぞれの CPU 性能値に、CPU 使用時間が比例するように計算する。CPU 性能値は、Office(Word, Excel, PowerPoint)の主要なユーザーアクションのレスポンス時間内の CPU 使用時間より、CPU の平均的な相対速度を求め、その相対速度から、Pentium III 500MHz の性能値を 100 とした場合の、それぞれの CPU の性能値を求める。

将来の PC 上での Disk の使用時間の予測も、CPU と同様に行う。なお、Disk 性能値は、IDE 5400rpm の Disk 性能値を 100 とした相対性能値である。

将来の PC 上での Network の使用時間の予測は、ユーザーアクションが、インターネット接続部分の送受信上限サイズに近い値で、一定時間送受信を行っている場合のみ、行う。将来の PC 上で、インターネット接続部分の送受信上限サイズが増加と共に、ユーザーアクションが送受信するサイズも増加すると予測し、送受信時間を予測する。

4.5.3. ユーザーアクション全体のレスポンス時間の予測

ResAnalyzer は、4.5.2で予測、将来の PC 上での、それぞれの CPU, Disk, Network 使用時間と、4.5.1で求めたそれぞれの使用時間ごとの依存関係をもとに、ユーザーアクション全体のレスポンス時間の PC 使用時間を再配置する。



図 6: 現在と将来の PC での、ユーザーアクションの PC 使用時間の配置

図 6 は、現在の PC のユーザーアクションの PC 使用時間の配置と、予測した将来の PC のユーザーアクションの PC 使用時間の例である。

4.6. ユーザーアクションに対する重要度を求める

ResAnalyzer は、ユーザーアクションが行われた時点でのアプリケーションの利用頻度データを求めることにより、ユーザーアクションごとの重要度を求める。アプリケーションの利用頻度データの構成を以下に示す。

- 現在起動しているアプリケーション i の起動時間 $T_{boot}(i)$
- 現在起動しているアプリケーション i のアクティブとなった時間 $T_{active}(i)$

これらのデータは、パフォーマンスデータのプロセスの起動時間と、ウィンドウメッセージの WM_ACTIVATE メッセージをフックした時間より求める。

ユーザーアクションの重要度は、ユーザーアクションが行われた時点での、ユーザーアクションを処理するアプリケーション i の利用頻度データより、式 2 で計算される値より求める。

$$w(i) = \frac{T_{active}(i)}{T_{boot}(i)} \quad \dots \text{式 2}$$

5. ResAnalyzer の検証

本節では、以下の点における ResAnalyzer の検証を行う。

- ユーザーアクションの認識精度の検証
- ユーザーアクションの予測精度の検証

これより、上記の ResAnalyzer の検証を行う。

5.1. ユーザーアクションの認識精度の検証

ユーザーアクションの認識精度の検証は、Office(Word, Excel, PowerPoint)の主要なユーザーアクションの認識精度について、目視による検証を行う。検証は、

- ユーザーアクション開始判定時刻
- ユーザーアクション終了判定時刻

に対し、行う。検証に利用した PC を表 1 に示す。

表 1: ユーザーアクションの認識精度の検証を行った PC 構

成

CPU	Pentium III 1.0G
Memory	SDRAM 512MB
HDD	IDE 5200rpm
Network	100BaseT - ADSL

認識対象となるユーザーアクションの一部を、以下に示す。なお、認識対象としたアクションは、全 50 種類である。

Word

- 起動
- 終了
- 書式変更
- 表示方法変更
- 数式オブジェクトの挿入

Excel

- 起動
- 終了
- グラフの作成
- マクロの実行

PowerPoint

- 起動
- 終了
- 背景テンプレートの適用
- オブジェクトの配置

連携アクション

- Excel のグラフを Word に貼り付け
- Word の図を PowerPoint に貼り付け

これらのユーザーアクションの認識結果を、表 2 に示す。

表 2: ユーザーアクションの認証精度の検証結果

対象アプリケーション	開始認識	終了認識
Word	100%	80%
Excel	100%	84%
PowerPoint	100%	90%
連携アクション	100%	100%

なお、検証中、ユーザーアクションを行っていないのに、ユーザーアクションを認識する場合が多数確認できた。

5.2. ユーザーアクションの予測精度の検証

ユーザーアクションの予測精度の検証は、5.1と同様、Office (Word, Excel, PowerPoint) の主要なユーザーアクションを、複数の PC 上で行い、それぞれの予測レスポンス時間と実機での測定レスポンス時間を比較することにより行う。予測のもととなる PC 構成を表 1 に示し、予測の対象となる PC 構成を表 3 に示す。比較には、ユーザーアクション中で行われた、CPU、Disk 読み込みの全使用時間の比較することで行う。また、ResAnalyzer のユーザーアクションの予測精度の検証結果を表 4 に示す。

表 3: ユーザーアクションの予測精度の検証を行った PC 構成

CPU	Pentium 4 2.8G
Memory	RDRAM 1024MB
HDD	IDE 7500rpm
Network	100BaseT - ADSL

表 4: ユーザーアクションの予測誤差

対象アプリケーション	CPU	DISK 読込
Word	0 ~ 20%	0 ~ 100%
Excel	0 ~ 20%	0 ~ 200%
PowerPoint	0 ~ 20%	0 ~ 120%
連携アクション	0 ~ 15%	0 ~ 150%

6. まとめ

本稿で提案した PC 性能評価手法は、従来の PC 性能評価手法では不可能であった、ユーザーの嗜好にあった PC 性能評価値を求めることが可能である。しかし、提案した PC 性能評価手法の実現には、(1) ユーザーアクションのレスポンス時間の認識、(2) ユーザーアクションのレスポンス時間の予測、(3) ユーザーアクションの重要度を測定、が必要であった。我々の開発した ResAnalyzer は、これらの問題点を解決し、提案する PC 性能評価手法を実現する。

ResAnalyzer は、ユーザーアクションをほぼ 80%以上認識することが可能である。しかし ResAnalyzer は、不要なユーザーアクションの認識も行っており、不要なユーザーアクションの絞り込みを行う必要がある。また、ResAnalyzer は、将来の PC 上でのユーザーアクションの CPU 時間を、20%程度の誤差で、予測することが可能である。しかし、Disk 読込時間に関しては、200%程度の誤差が発生する。これは、フラグメントの影響が考えられる。

今後、Network の予測検証、Disk のフラグメントの影響の予測を行う。また、ファイルキャッシュや、ページングの影響の予測も行っていく予定である。

参考文献

- [1] <http://www.specbench.org/>
- [2] David Bailey, Tim Harris, William Saphir, Rob van der Wijngaart, Alex Woo, and Maurice Yarrow: "The NAS Parallel Benchmarks 2.0," NASA Ames Research Center Report, NAS-05-020, 1995.
- [3] <http://www.sisoftware.net/>
- [4] Ziff-Davis Corp, "Labs Notes: Benchmark 97: Inside PC Labs' Latest Tests", PC Magazine Online, Vol. 15, No. 21, December 3, 1996.
- [5] <http://www.bapco.com/>
- [6] <http://www.sisoftware.net/>
- [7] <http://www.futuremark.com/>
- [8] <http://www.veritest.com/benchmarks/winbench/winbench.asp>