

Technology Reports

FOMA 905i 搭載アプリケーション機能

FOMA 905i シリーズにおける新規アプリケーション機能として、ソフトウェア自動更新機能、Flash[®]*1 機能追加、ミュージックプレーヤ利用履歴送信機能、分散型音声認識機能を開発した。

はまだ てつや とべ あきこ
浜田 哲也 戸部 章子
いちのせ あきひろ とりもと なおみ
一瀬 晃弘 鳥本 尚実

1. まえがき

移動端末は、ユーザが今まで以上に「安心・安全」に利用することができ、かつ「先進的な機能を使いやすいサービス」として常に提供していく必要がある。そのため FOMA 905i シリーズでは、「安心・安全」に利用するために「ソフトウェア自動更新機能」を開発し、「先進的な機能を使いやすく」利用するために「Flash 機能追加」、「分散型音声認識」による機能拡充を図った。

また、ユーザ許諾の下、ユーザが利用した楽曲の履歴を音楽配信サービス提供者（CP：Contents Provider）に送信することで、CP がマーケティングへの活用を図れるように「ミュージックプレーヤ利用履歴送信機能」を開発した。

本稿では、これら 4 つの新機能の概要について解説する。

2. ソフトウェア自動更新機能

FOMA 905i シリーズより搭載したソフトウェア自動更新機能について述べる。

2.1 導入背景

ソフトウェア更新機能は、移動端末のソフトウェアへの修正が必要な場合に、ユーザの操作により不具合を修正する機能である。

近年、移動端末のソフトウェア開発では開発プロセスの改善など、さまざまな高品質化施策を実施しているが、高機能化に伴う開発規模の増大により市場投入後に不具合が判明する事例が発生している。

このような背景の下、ユーザが少しでも早く、修正されたソフトウェアを使用できるよう、FOMA 900i シリーズから、無線によるソフトウェアの更新機能を搭載し、さらに FOMA 904i シリーズでは更新が必要なユーザにアイコンで通知する「更新お知らせ機能」を搭載した。しかし、通知されたとしても更新するにはユーザの手動操作が必要だった。このため、ユーザの操作を必要としないソフトウェア自動更新機能の実現が課題であった。

2.2 導入の課題と対策

ソフトウェア自動更新機能の導入

時の 2 つの課題とその対策について以下に述べる。

課題① ソフトウェア書換え中は移動端末のすべての機能が利用できない。

課題①より、ソフトウェア自動更新は、ユーザが移動端末を使用しない時間帯に実行されることが望ましいが、ユーザごとにその時間帯は異なっている。一方、ネットワークの負荷を低減するためには、修正用ファイルのダウンロードが同じ時間帯に集中しないよう平準化する必要がある。

このため、修正用ファイルのダウンロード処理と、ソフトウェア書換え処理とを分離し、ダウンロード処理の開始契機はネットワーク側の制御で行い、ソフトウェア書換え開始契機は移動端末でユーザが設定可能とする機能を開発した。

課題② ソフトウェア書換え中に電池が外れると書換えに失敗する。

本機能はユーザの移動端末ソフトウェアを自動で更新するため、書換えに失敗し、移動端末が使用できなくなることは絶対にあってはならない。

*1 Flash[®]：音声やベクタグラフィックスのアニメーションを組み合わせたコンテンツを作成するソフトウェア、または作成されたコンテンツのこと。
Flash は、Adobe Systems Inc. の米国およびその他の国における商標または登録商標。

そこで、書換え中に電池が外れたとしても、再度電源を供給すれば途中から書換えを再開できるフェイルセーフ機能を搭載した。

2.3 自動更新手順

まず、自動更新の有効/無効を切り替えるための自動更新設定について述べる。自動更新設定の画面例を図1に示す。ユーザは自動更新設定により、自動更新を行う「自動で更新」、更新通知を表示するだけで自動更新は行わない「更新の通知のみ」、自動更新も更新通知表示も行わない「設定しない」を選択することができる。また、ユーザの都合により書換え開始の曜日、時刻の変更も可能である。初期設定値は、より多くのユーザに自動更新を利用してもらうために、「自動で更新」としている。

次に、自動更新処理の流れを述べる。本機能では以下の3つのステップで自動更新を実現している。

- ①ダウンロード予約設定
- ②修正用ファイルのダウンロード
- ③ソフトウェアの書換え

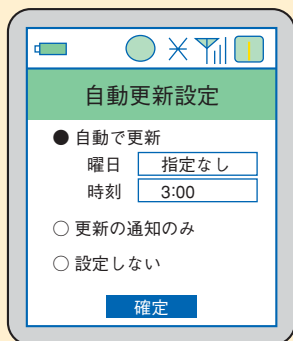


図1 自動更新設定の画面例

ソフトウェア更新用サーバからの信号を受信した移動端末は、更新の必要可否を判断し、更新の必要がある場合は予約設定を行う。予約設定とは、修正用ファイルのダウンロード日時をサーバに問い合わせ、ダウンロード時刻を取得する機能である。この機能により、サーバへのトラフィックを分散することが可能となる。

移動端末は、予約設定で取得した日時に自動でソフトウェア更新機能を起動して、修正用ファイルのダウンロードを行う。ダウンロード完了後は、書換え開始時刻を通知するアイコンを待受画面上に表示する。

その後、移動端末は自動更新設定画面で指定した書換え時刻に自動で書換え処理を開始する。なお、他の機能動作中などで書換えが開始できない場合は、次のタイミング（初期設定は翌日の午前3時）に再度書換えを開始する。以後、書換えが完了するまでこの動作を繰り返す。

3. Flash 機能追加

3.1 導入背景

Flashはmova 505iシリーズからの搭載機能であり、i-modeブラウザ上での表示、iチャンネルサービスでの利用、待受画面やメニューなどの利用、多岐に利用されている。しかし、前のページに戻ったときに、直前のフォーカス位置に戻らない、また、一律再配布不可のためメール添付ができないなどの制限があった。

そこで、ユーザやコンテンツ作成者の利便性向上、表現力の向上のた

めに、機能追加を行った。3.2章と3.3章にて、追加された主な2つの機能について述べる。

3.2 データストレージ機能

ファイルの指定により、ファイル情報の記憶領域（データストレージ領域）を作成・利用する機能である。領域のサイズはコンテンツ作成者が指定可能であり、最大10KBの領域を確保することが可能である。また、1つのFlashファイルに対し、1つのデータストレージ領域が割り当てられ、IDで関連付けることにより、Flashファイルが他のデータストレージ領域を使用できないようにしている（図2）。データストレージ機能の使用例として、ここではブラウジング時のフォーカス位置保存について述べる。

Webページの画面遷移では、ブラウザアプリケーションが前ページでフォーカスが当たっていた部分を保

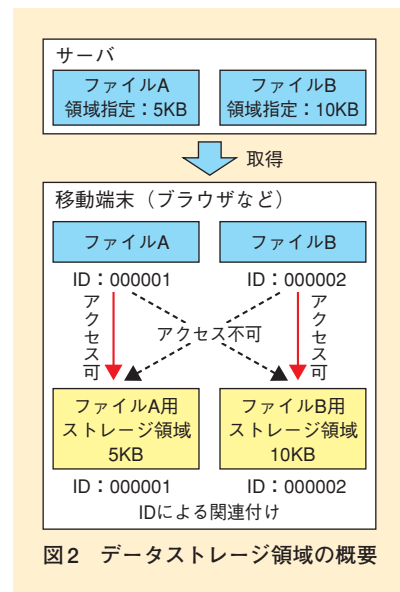


図2 データストレージ領域の概要

持っているため、リンク押下などによりページ遷移をしたのち前のページに戻った際に、以前フォーカスが当たっていた位置にフォーカスが戻る仕組みとなっている。しかし、Flashコンテンツでは、ブラウザアプリケーションがフォーカス位置を取得する手段がなく、前ページに戻ると一番上にフォーカスが戻ってしまう動作となっていた。そこで、本データストレージ領域に直前のフォーカス位置を保持しておくことで、Webページと同様に前ページに戻っても元の位置にフォーカスを表示することを可能とした。

データストレージ領域はコンテンツ作成者が自由に利用できるため、フォーカス位置保存以外にも、Flashコンテンツ上でユーザごとに異なる文字列を表示するなど、これまで実現できなかったFlashコンテンツでのサービスが可能となる。

3.3 メタデータ対応

以前のFlashファイルは、ファイルを構成するために必要なデータのみを保持していたため、ファイルに対し補足情報を付与することができなかった。

新たにFlashファイル内にメタデータ^{*2}領域を作成し、以下の3点の情報を付加した。

①再配布不可識別子

現状では、Flashコンテンツは一律再配布不可コンテンツとして扱っており、コンテンツ作成者による指定ができない。今回メタデータ領域内に再配布の可否を指定

する領域を新設した。これにより、コンテンツ作成者は外部出力可否の制御が可能となり、Flashコンテンツの外部出力（メール添付、外部メモリ保存など）を伴う新しいサービスの提供を可能とした。

②背景画像透過の有無

Flashの再生領域内でも背景部分を透過表示することを可能とした。これにより、Webページの背景部分（背景色、静止画像のみの背景画像）が表示され、矩形以外の画像の表示が可能となり、より表現力のあるWebページをユーザが閲覧可能となる。

③絵文字デフォルト色指定

Flashファイルにおいて、文字色はコンテンツ作成時に色指定をする必要がある。その際、絵文字を入力しても絵文字のデフォルト色ではなくオーサリングツール^{*3}での指定色で表示されるため、コンテンツ作成者はそれぞれ絵文字の色を指定する必要があった。

本指定をすることで、絵文字は指定色ではなく絵文字のデフォルト色で表示されるようになるため、Flashコンテンツ作成の手間を省くことができる。

4. ミュージックプレーヤ 利用履歴送信機能

4.1 導入背景

移動端末のミュージックプレーヤ機能を使って楽曲視聴を楽しむ利用者が増加している。ユーザが移動端末で聴く楽曲は、音楽配信サービスにより移動端末自体でダウンロード

した楽曲に加え、CDなどからリッピング^{*4}した楽曲も多く利用されている。CPは、CDなどからリッピングした楽曲を含めて、ユーザが聴いている楽曲のジャンル、アーティスト名などの情報と関連付けを行うことで、ユーザごとの嗜好に合わせた楽曲の提示が可能となり、楽曲のオンライン販売の促進が見込まれる。また、ユーザも自分の好みの楽曲を簡単に見つけることが可能になる。

そこで、ユーザ個人の嗜好に合った楽曲や多くの人に聴かれているランキングの提供など音楽配信サービス向上のため、ユーザが実際に楽曲を聴いた履歴情報をCPに送信可能とするミュージックプレーヤ利用履歴送信機能（以下、利用履歴送信機能）を開発した。本機能により、楽曲配信サービスなどのWebページ上で、ユーザが利用履歴送信タグ^{*5}を選択した際に、利用履歴情報を当該サーバに送信することが可能となる。

4.2 利用履歴送信手順

利用履歴送信機能の動作フローを図3に示す。

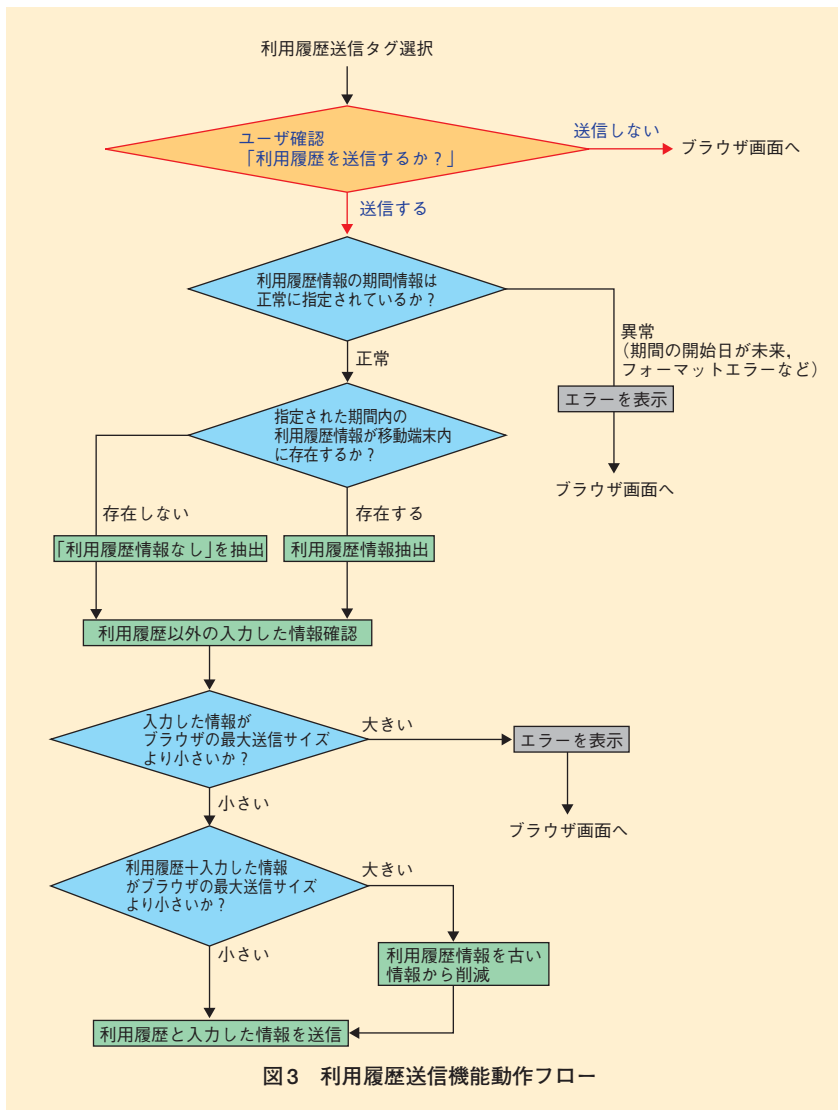
ユーザが利用履歴送信を要求するタグを選択すると、ミュージックプレーヤの利用履歴を送信するか否かを選択する画面が表示される。ここで、ユーザが「利用履歴を送信しない」を選択した場合は、利用履歴情報を送信せず、ブラウザ画面に戻る。ユーザが「利用履歴を送信する」を選択した場合には、Webページ上のタグで指定されている期間内に該当する移動端末内部で管理している

*2 **メタデータ**：コンテンツデータそのものではなく、そのコンテンツデータに関連する情報。例えば、画像ファイルに対する撮影場所の位置（緯度経度）情報などが該当する。

*3 **オーサリングツール**：文字や画像といった異なる種類のデータを編集して、1つのソフトウェアやコンテンツを作るためのアプリケーションソフトウェア。

*4 **リッピング**：音楽CDなどに記録されているデジタルデータを、PCで扱いやすいデータ形式に変換すること。

*5 **タグ**：Webページ内で見出しやリンクなどを指示する記述方法。



利用履歴情報をブラウザに通知する。

ブラウザでは、利用履歴以外にユーザがWebページ上で入力した情報（例：アンケートなどでコメント欄へユーザが入力した情報）などと一緒に利用履歴を送信する場合も含めて、ブラウザで送信できる最大サイズ内に、入力した情報と利用履歴情報が収まるように、利用履歴の古い方から順に情報を削減したうえで、CPに利用履歴を送信する。

5. 分散型音声認識機能

5.1 導入背景

音声認識機能は、音声による移動端末への入力手段を提供する機能である。例えば、ユーザの発話内容を文字へ変換することで、文字入力に利用できる。

音声認識機能では、入力された音声から対応する文字を認識するための辞書が必要となる。数百程度の単

語を識別するならば、認識用の辞書を移動端末内に保持できるが、短文を変換するような数万以上の単語を識別する認識を行うためには、辞書の大きさが膨大となり、認識処理をサーバで行う必要がある。移動端末だけではなく、サーバも利用する音声認識を分散型音声認識（DSR：Distributed Speech Recognition）と呼ぶ。

FOMA 905iシリーズではiアプリの新たな入力手段としてDSR機能を利用できるように以下の開発を行った。

- ・移動端末での音声特徴量抽出処理
- ・移動端末とサーバとを連携させるための通信プロトコルの新規規定
- ・iアプリから利用するためのAPI（Application Program Interface）^{*6}

これにより、例えば旅行会話日英／英日翻訳iアプリでは、これまで旅行会話文をキー入力していたところを、音声入力が可能となり、利便性を向上させることができる。

移動端末内の構成を図4に示し、各要素については、それぞれ5.2～5.4章にて述べる。

5.2 音声特徴量の抽出処理

音声認識では、ユーザの音声以外の音（雑音）を除去する処理と、ユーザの音声から認識に有効な情報を算出した音声特徴量を抽出する処理とが必要となる。この処理部分をフロントエンド（FE：Front-End）と

*6 API：OSやミドルウェアなどが提供する機能を、他のソフトウェアが利用するためのインタフェース。

呼び、FOMA 905iシリーズにて新たに搭載した。

搭載されたFEでは、音声特徴量を抽出するほかに、ユーザの音声の有無/音声の大きさ/音声に対する雑音の大きさを算出し、iアプリに通知する。iアプリではこれらの情報から「もっと大きな声で話してください」などのアドバイスをユーザに提示可能となる。また、ユーザの1回の発話で最大30秒間の音声特徴量を保持できる。発話時間が30秒を超えると、音声入力が自動的にFEにて中断される。30秒以内ならば、ユーザの指示やiアプリの判断で発話を中断することもできる。FEでは中断された時点までの音声特徴量を抽出し、iアプリへ音声特徴量の送信を行う。

5.3 通信プロトコルの新規規定

移動端末と認識用のサーバとの間で通信を行うためのプロトコルを

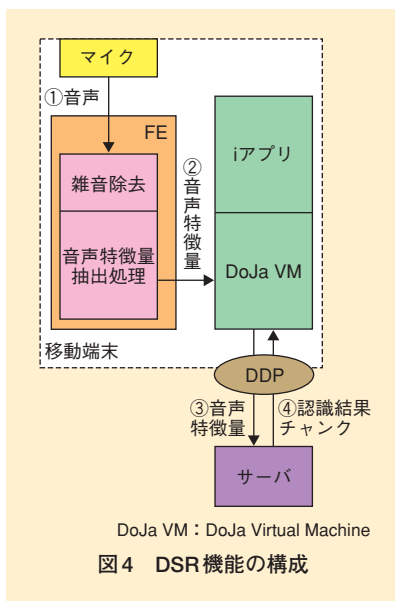


図4 DSR機能の構成

DDP (DoCoMo DSR Protocol) として新たに策定した。

DDPは、HTTP上で動作するプロトコルであり、音声認識処理の進捗に応じた移動端末からの命令とサーバからの応答を規定している。また、音声特徴量を格納する出力フォーマットも規定しており、ETSI (European Telecommunications Standards Institute)*7 ES 202-050 [1]準拠の出力フォーマットにて構成される。

5.4 iアプリからの利用

DSR機能を利用するためにiアプリのAPIを新設した。API利用シーケンスの基本的な例としてユーザ操作で発話を終了する場合を図5に示す。

まず、iアプリはFEからFE自身の情報を取得する(図5①)。次に、取得したFEの情報をを用いて、サーバとのネゴシエーション*8を行う(図5②)。ネゴシエーション結果を基に、FEに音声特徴量抽出処理の開始を指示する(図5③)。iアプリ

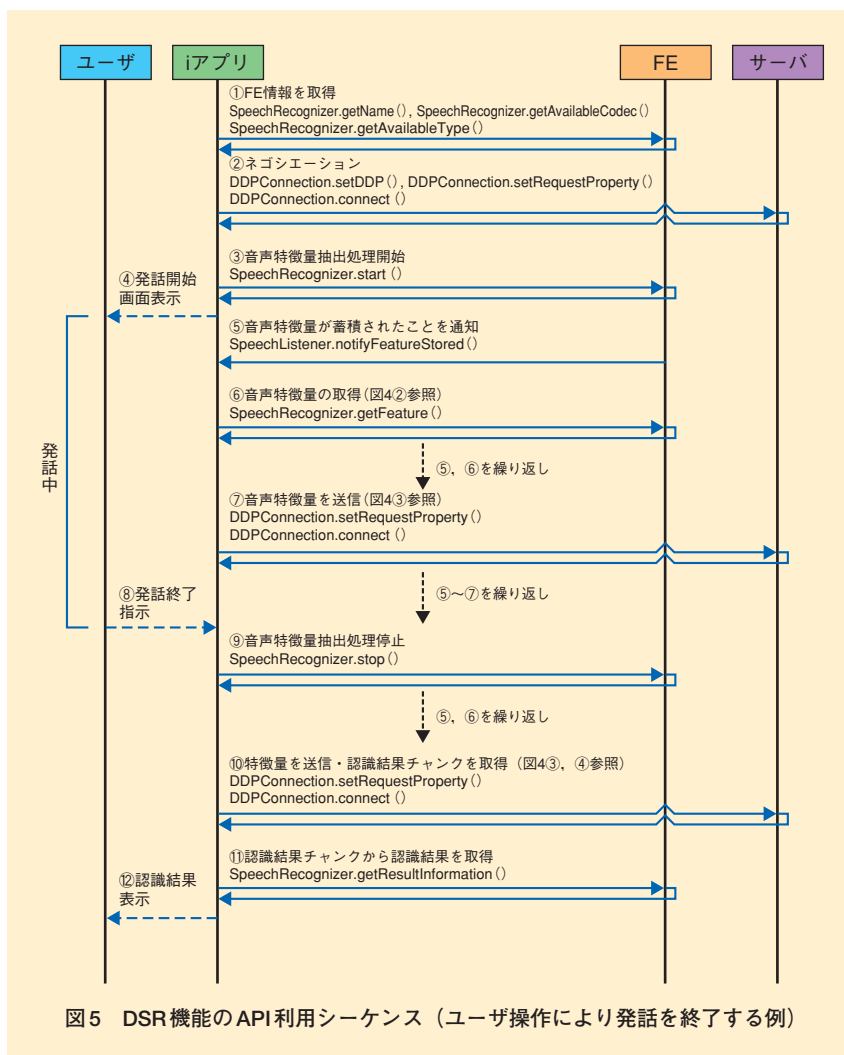


図5 DSR機能のAPI利用シーケンス (ユーザ操作により発話を終了する例)

*7 ETSI: 欧州電気通信標準化機構。ヨーロッパの標準化団体。電気通信技術に関する標準化を行っている。本部はフランスのSophiaAntipolisにある。

*8 ネゴシエーション: 移動端末とサーバ間で認識処理に利用する条件を一致させる確認処理。

は発話開始画面などを表示してユーザに発話開始を促す(図5④)。ユーザの発話により入力された音声からFEにて音声特徴量が抽出され、蓄積されるとiアプリにその旨が通知される(図5⑤)。iアプリはFEから音声特徴量を取得し(図5⑥)、サーバに送信する(図5⑦)。ユーザが発話終了を指示する(図5⑧)まで、これらの処理(図5⑤～⑦)を繰り返す。iアプリは音声特徴量抽出処理停止をFEに指示し(図5⑨)、停止指示前に入力された音声から抽出された音声特徴量をすべてiアプリ

がサーバに送信すると、サーバから認識結果チャンク*⁹が出力される(図5⑩)。その認識結果チャンクをFEで解析し、認識結果を取得する(図5⑪)。以上により、iアプリは認識結果をユーザに提示することができる(図5⑫)。

6. あとがき

ソフトウェア自動更新機能、Flash機能追加、ミュージックプレーヤ利用履歴送信機能、分散型音声認識機能をFOMA 905iシリーズでの新規機能として搭載することによ

り、移動端末の更新機能や表現力、入力手段の向上を図った。

今後も、搭載している機能を使いやすく拡張し、入出力機能を高度化することを目指し、ユーザが求める移動端末の開発を進めていく予定である。

文献

- [1] ETSI: "Speech Processing, Transmission and Quality Aspects (STQ); Distributed speech recognition; Advanced front-end feature extraction algorithm; Compression algorithms," ES 202-050 v1. 1.5, Jan. 2007.

*9 認識結果チャンク: DDPで規定されたフォーマットに従って記述された認識結果。認識された文字情報そのもの以外に、認識結果の候補数や信頼値が含まれる。