

超大規模バイオメトリック 認証システムとその実現

蘇雷明・坂本 静生

要 旨

インド国民ID制度は、世界人口の約1/6にあたる12億人をバイオメトリック認証で識別しようとする、これまで世界に類をみない超大規模システムです。他にもこのような国家的な規模の認証システムの検討が進んでおり、NECは実現へ向けて開発を進めています。本稿ではこの超大規模バイオメトリクス情報を処理するシステムについて説明します。

キーワード

- バイオメトリック認証
- マルチモーダル
- インド国民ID制度
- 大規模データベース
- スケーラビリティ

1. はじめに

人が誰であるかを確認・認証する処理は、社会生活を送るさまざまな場面で必要不可欠です。特に2001年9月11日に起きた米国同時多発テロは、その重要性を大きく変えました。しかし、カードの所持や、ID・パスワードの利用などによる本人の証明は、漏えい・偽造・なりすましを本質的に防ぐことができません。そのため、本人のバイオメトリック情報を用いるバイオメトリック認証技術が脚光を浴びることとなり、パスポートや出入国管理における応用が加速的に進行了しました。

国民IDも、バイオメトリクスの応用が急速に浸透する分野の1つです。NECはこれまでも、米国国立標準技術研究所 (National Institute of Standards and Technology : NIST) による第三者評価により、世界トップの認証精度が証明された指紋認証を利用する南アフリカの国民IDなどのビジネスを進めています¹⁾。しかし近年プロジェクトが始まったインド国民ID制度 (以下、インドUID) は、最終的に世界人口の1/6である12億人をバイオメトリクスで認証し、インド電子政府の認証基盤とするこれまでにない超大規模なものです。弊社はこのシステムの実現へ向け、数々の新しい技術的課題に取り組んでいます。

本稿ではインドUIDに代表される、超大規模なバイオメトリック認証システムに対する弊社の取り組みについて紹介します。

2. インドUID

インドUID²⁾ はインド電子政府の認証基盤であり、納税、福祉、銀行などの認証業務まで、幅広い応用が期待されています。このIDの業務発行はUIDAI (Unique Identification Authority of India) と呼ばれるインド政府の政庁が担当し、今後数年間の歳月を費やしてインド全国民一人ひとりに固有のIDを発行していく予定です。

さまざまな認証業務に使うためには、各人唯一のIDを割り当て、重複したIDが発行されていないことが重要であり、顔と10指の指紋及び両目の虹彩を含むバイオメトリック情報を採集して、登録済みの情報との重複がないかチェックすることで保証します。プロジェクトは2009年に着手しており、2012年内に2億人の登録照合を完了する見通しですが、これ以降はID発行のスピードを更に加速して、2014年末までに6億人の登録を完了する予定です³⁾。

バイオメトリック情報による重複チェック業務は、バイオメトリック情報の採集管理と、バイオメトリック認証による登録・照合に分けられます。このバイオメトリック認証システムの全体規模はこれまで世界が経験したことのないものであり、弊社は登録・照合業務を担当し、処理を続けています。

3. 超大規模バイオメトリック認証システムの課題

インドUIDに代表される超大規模バイオメトリック認証シ

システムを実現するためには、さまざまな課題を解決しなければなりません。弊社はシステムを分析・検討することにより、次の4つの課題を同定しました。

課題1：マルチモーダル認証精度

課題2：マルチモーダル認証処理最適化

課題3：スケーラビリティ

課題4：高可用性環境の構築

以下で各課題を説明します。

3.1 課題1：マルチモーダル認証精度

高い認証精度を達成するためには、認証に向く品質でバイオメトリック情報を採集することが最も重要です。しかしインドUIDのような超大規模バイオメトリック認証システムでは、採集自体が大きな課題となり得ます。インドUIDは、顔・両目の虹彩・両手全指の指紋、これら3つのモダリティを採取することになっていますが、すべてを可能な限り高い品質で採集するためには以下を考慮する必要があります。

まずインドは国土が非常に広く、世界で2番目に多い人口を抱えています。また、システム稼働までの時間を短縮するために、多くの採集要員を投入しています。しかし、採集者の経験や受けたトレーニングの差異があり、バイオメトリック情報の品質が採集者に依存する可能性を排除できません。また、バイオメトリック情報の採集に使われる多数のデバイスが複数の会社から調達されていることも、品質が安定しない要因となり得ます。

インドUIDは国民IDなので、15歳以上の全国民をカバーしなければなりません。人々の住む環境、職種などの差異に起因する品質も問題となります。例えば、農業を中心とした肉体労働者の一部は指紋の損傷が激しいなど、広い国土でさまざまな文化・業種に携わる方による傾向が異なることが分かっています。また、ひげを生やす習慣などもバイオメトリック認証にとっては問題となります。すなわちデバイスなどの採集条件、採集者、採集対象のいずれも認証精度に影響する可能性があります。

バイオメトリック認証システムを提供する弊社としては、上記の考えうる劣化要因を考慮したうえで、採集されるバイオメトリック情報を、認証に向くデータとしてできるだけ多く受け入れられるような照合アルゴリズムを開発しなければなりません。

3.2 課題2：マルチモーダル認証処理最適化

一般に情報処理分野では、登録した母集合からあらかじめ検索のための索引を作成しておくことにより、対応する個体を高速に検索する手法を用います。しかしバイオメトリクスの場合には、経年変化や採取時の変動などのために、同一人物から採取したバイオメトリック情報であっても厳密に一致することはありえません。このため一般に索引を作成・利用することは難しく、例えば登録希望者が重複していないことを確かめるには、登録済みの全データに対して照合を行うしかありません。特にインドUIDでは厳密な重複チェックのため、氏名、性別、年齢などの身上情報を利用してあらかじめ絞り込むことが禁じられており、計算量が膨大に膨れ上がります。

例えばインド国民12億人全員の重複チェックには、約 7.2×10^{17} 回のマッチングが必要となります。2014年末に予定している6億人をチェックする場合でも約 1.8×10^{17} 回に上ります。実際に6億人の重複チェックを3年以内以内実施するためには、無停止で1秒あたり19億回以上のマッチングが可能なシステムを実現しなければなりません。

したがって、虹彩や指紋それぞれを個別にすべて照合して結果を融合するのではなく、最適な利用を行わなければなりません。このマルチモーダルの最適化を行うためにはそれぞれのバイオメトリクス技術を熟知していなければならず、実験レベルではない運用システム上で実現できるのは世界的にも数社を数えるのみです。

3.3 課題3：スケーラビリティ

個々のバイオメトリック認証処理を最適化できたとしても、次々と入る照合要求に対して、超大規模バイオメトリックデータベースからの照合結果を効率的に応答することができなければ、スケーラビリティに問題が生じます。

バイオメトリック認証のトランザクション処理の模式図を図1に示します。照合の要求がシステムに入力されると、内部の負荷・処理状況により登録バイオメトリック情報の配置計画・サーバ上での実行計画を作成した後、各サーバへ照合要求を配信して照合処理を実施、照合結果を回収して最終的な応答を行います。これらの管理コストにわずかでも遅延が生ずると、照合サーバにアイドリング期間が発生することになり、照合効率の低下や、照合サーバを追加してもトランザ

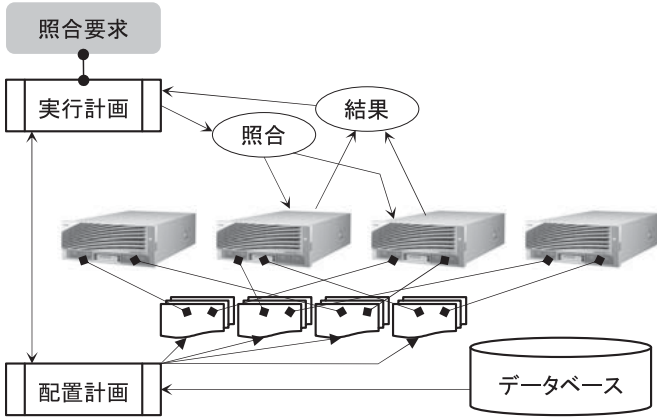


図1 分散処理構成のイメージ図

クッション処理のスループットが上昇しないスケーラビリティ限界に到達してしまうことになります。

3.4 課題4：高可用性環境の構築

システムを構成するサーバの台数が非常に多いこと、特に照合処理を担当するサーバはCPU使用率が100%に近い状態で酷使することから、一般に障害発生率は高くなります。しかし、政府の認証基盤として利用されるシステムなので厳しい耐障害性能が要求されており、以下のような実現への大きな課題があります。

- 1) 障害が起きても、バイOMETリック情報を損失しないこと
- 2) 単一機材の障害であれば、すべての業務が継続可能であること
- 3) 複数データセンターでの運用により災害対策を講ずると同時に負荷分散も実現すること

4. 超大規模バイOMETリック認証システムの実現

前述の課題のうち、実装の課題である4を除く、課題1から3に関する弊社の解決方法を紹介します。

解決1：マルチモーダル認証精度

一般に指紋、虹彩、顔といったバイOMETリック情報は、人種や習慣、年齢などへの依存性があります。例えば、人種による顔の違いや、手の大きさ、指の長さなどに特徴があることが直感的に理解できるかと思います。

弊社のシステムでは、それぞれのモダリティに対応する認証アルゴリズムに、各案件の状況に合わせたチューニングが可能です。実際にインドUIDでは、インド国民の実情に合わせた最適なチューニングを施しています。例えば、ターバンやひげなどがある顔画像の認証をより高精度に行うために、インドの方々のデータによってアルゴリズムを強化して最適なパラメータを選択しました。

指紋では、若く成長期にあたる方から、老年期に至る指紋の隆線認識が非常に困難な方、また、農業など肉体力労働者のように指紋にダメージが多い方を考慮しました。また、インドの方々の指の長さは比較的長いことなど人種の特徴も加味したパラメータを、評価検証を繰り返すことで高い認証精度を実現しています。

解決2：マルチモーダル認証処理最適化

認証精度を保ちながら高速な照合処理を実現するには、すべての指紋・両目の虹彩・顔すべてのバイOMETリック情報を照合するのではなく、必要最小限で最適な組み合わせのみを実行します。ここでは便宜上、2つのバイOMETリック情報A・Bだけを用いる簡略化した仮想的なモデルで説明します。

バイOMETリック認証の精度は、セキュリティの指標となるFAR (False Acceptance Rate：他人を間違っ受ける率) と、ユーザビリティの指標となるFRR (False Rejection Rate：本人を間違っ拒絶する率) で計測します。

ここでは、バイOMETリック情報A単独の認証精度を(FAR_a, FRR_a)・単一CPUコアにて1秒で処理できる件数をS_aとします。同じくバイOMETリック情報Bでは(FAR_b, FRR_b)・S_bとなります。2つのバイOMETリック情報間で相関がないと仮定することで、さまざまな組み合わせ方法での認証精度と処理能力が予測できます。

・ 組み合わせ (1)

どちらかのバイOMETリック認証結果が、本人と判断する閾値を満たす場合に、本人と判断 (OR)

$$\text{毎コア毎秒処理件数} : S_1 = 1 / (1/S_a + 1/S_b)$$

$$\text{本人誤拒絶率} : FRR_1 = FRR_a \times FRR_b$$

$$\text{他人誤受入率} : FAR_1 = 1 - (1 - FAR_a) \times (1 - FAR_b)$$

この組み合わせでは毎コア毎秒処理件数とFARが劣化しますが、FRRが向上します。

・ 組み合わせ (2)

両方のバイOMETリック認証結果とともに、本人と判断する閾値を満たす場合に限り、本人と判断 (AND)

毎コア毎秒処理件数： $S_2 = 1 / (1/S_a + 1/S_b)$

本人誤拒絶率： $FRR_2 = 1 - (1 - FRR_a) \times (1 - FRR_b)$

他人誤受入率： $FAR_2 = FAR_a \times FAR_b$

よって毎コア毎秒処理件数は組み合わせ (1) と同一ですが、FRRが劣化しFARが向上します。

・ **組み合わせ (3)**

バイOMETリック情報Aの認証結果が、本人と判断する閾値を満たす場合に限り、バイOMETリック情報Bにより認証して閾値で本人と判断

毎コア毎秒処理件数： $S_3 = 1 / (1/S_a + P_3/S_b)$

ただし、 P_3 は1つ目の認証結果が閾値を満たす確率

本人誤拒絶率： $FRR_3 = FRR_a + FRR_b$

他人誤受入率： $FAR_3 = FAR_a \times FAR_b$

最初の2つと異なり、この組み合わせではシステムの要求認証精度を満たすようFAR及びFRRを調整することにより、毎コア毎秒処理件数を向上させることができます。特に処理件数能力の差が大きなバイOMETリック認証を組み合わせる場合、適切な閾値を選択することにより高い効果が得られます。例えば $S_a = 100$ 万件・ $S_b = 5$ 万件・ $P_3 = 10\%$ のとき、組み合わせ (3) は組み合わせ (1) に比較して約7倍の処理速度が得られます。

弊社がインドUIDプロジェクトに提供したシステムは、指紋・虹彩・顔を用いる多段階照合方式の処理フローを採用し、認証精度を保ちながら処理件数能力の最適化を行っています。この多段階照合方式は、部分的にAND/OR条件なども多様に結合した構成を取っており、更に、将来的に起こり得るインドUIDの採集バイOMETリックデータの品質変化にも配慮した拡張性や可用性を備えています。

このマルチモーダル認証の最適化を可能にしたのは、弊社の先端的な指紋・虹彩・顔の各バイOMETリック認証技術研究開発成果^{4) 5)} 並びに、インドUIDを対象とする評価試験へかたむけた努力の結晶です。

解決3：スケーラビリティ

超大規模データベースにまで対応可能なスケーラビリティを確保するためには、照合サーバ群の制御ロジックにスケールアウト・規模透過性を持たせる必要があります、技術的にも最大の難関と言えます。一般的に、1台のサーバが処理できる要件を1単元としたとき、10台のサーバで並列に処理しても全体の処理能力は10単元にはなりません。この傾向はサーバ台数

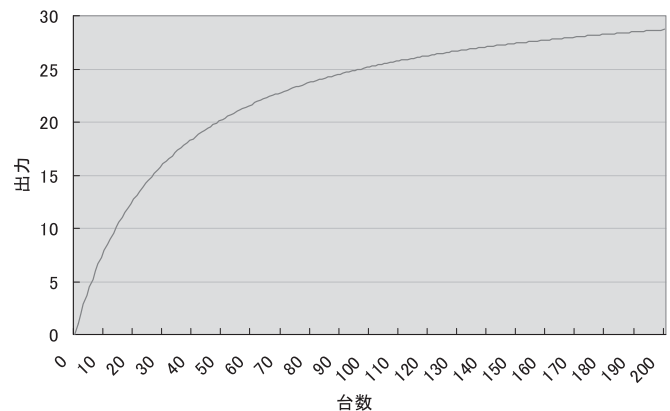


図2 スケーラビリティの限界 (F=97%のとき)

が増えるにつれて顕著となり、処理能力の向上率はどんどん劣化していきます。

劣化率がサーバ台数の増加率に等しくなったとき、それ以上サーバを追加しても処理能力は停滞します。各サーバ間の相関影響要素による劣化を除いた実効性能の割合をFとしたとき、n台のサーバによる並列処理の実効出力は $1/(1-F+F/n)$ で表せます。図2にFが97%の場合のグラフを示します。相関影響要素が3%の場合、サーバをいくら追加しても実行出力は30単元を超えることができません。例えば1,000台の照合サーバを並列化したとき、実行出力900単元・実効率90%を達成するには、相関影響要素を0.011%以下に抑える必要があります。

弊社のシステムでは、次の手法でサーバの管理を行っています。

- ・ 内部的なジョブの運搬をバッチで処理
- ・ トランザクションを複数フェーズで状態管理

トランザクションを数百台レベルの照合サーバで処理するとき、個別に実行計画・配置計画・デリバリ・実行結果の集約を行うことは管理コストが非常に高くなります。そのため一定期間内のトランザクションを1つの実行計画で管理することで、サーバ間交信の回数、分散管理の計算に使うCPUパワーを数百分の一まで低減しました。

またトランザクションは、1つのトランザクションとして管理しようとする状態・進行監視のコストが高く、内部処理の遅延に遅延が発生しやすくなります。そのために内部的に複数のフェーズへ分割してきめ細やかな管理を行うことで、実行効率の向上を図っています。

5. おわりに

バイOMETリック 認証技術は、指紋に続いて顔や虹彩なども日進月歩成熟を続けており、実用化されています。更に、バイOMETリクス 認証技術を国民IDなどの国家規模の認証システムに用いることが一般的になりつつあります。弊社はバイOMETリック 認証分野において、長年蓄積した技術に加えて超大規模かつ高信頼なシステム化技術を生かし、これからも日本及び世界の国々がより安心・安全な社会になっていくために貢献していきます。

参考文献

- 1) 坂本静生：“バイOMETリクス製品とソリューションの現状と展望,” NEC技報 Vol.63 No.3, pp.14-17, 2010.9
- 2) Unique identification Authority of India(UIDAI)
<http://uidai.gov.in/>
- 3) The TIMES of India, Nov 14th 2009 / The Economist, Nov 17th 2011.
- 4) 溝口正典,原雅範：“指紋掌紋の照合技術,” NEC技報 Vol.63 No.3, pp. 18-21, 2010.9
- 5) 今岡仁,早坂昭裕ほか：“顔認証技術とその応用,” NEC技報 Vol.63 No.3, pp.26-30, 2010.9

執筆者プロフィール

蘇 雷明

官公ソリューション事業本部
第二官公ソリューション事業部
マネージャー

坂本 静生

官公ソリューション事業本部
第二官公ソリューション事業部
上席アドバンストテクノロジスト
(ソフトウェア)
エグゼクティブエキスパート

NEC 技報のご案内

NEC 技報の論文をご覧くださいありがとうございます。
ご興味がありましたら、関連する他の論文もご一読ください。

NEC 技報 WEB サイトはこちら

NEC 技報 (日本語)

NEC Technical Journal (英語)

Vol.65 No.2 ビッグデータ活用を支える 基盤技術・ソリューション特集

ビッグデータ活用を支える基盤技術・ソリューション特集によって
ビッグデータを価値に変える NEC の IT インフラ

◇ 特集論文

データ管理 / 処理基盤

超高速データ分析プラットフォーム [InfoFrame DWH Appliance]
SDN 技術で通信フローを制御する [UNIVERGE PF シリーズ]
大量データをリアルタイムに処理する [InfoFrame Table Access Method]
大量データを高速に処理する [InfoFrame DataBooster]
ビッグデータの活用最適なスケールアウト型新データベース [InfoFrame Relational Store]
高い信頼性と拡張性を実現した Express5800 / スケーラブル HA サーバ
大規模データ処理に対する OSS Hadoop の活用
大容量・高信頼グリッドストレージ iStorage HS シリーズ (HYDRAStor)

データ分析基盤

ファイルサーバのデータ整理・活用を支援する [Information Assessment System]
超大規模バイオメトリック認証システムとその実現
WebSAM の分析技術と応用例～インバリエント分析の特長と適用領域～

データ収集基盤

スマートな社会を実現する M2M とビッグデータ
微小な振動を検知する超高感度振動センサ技術開発とその応用

ビッグデータ処理を支える先進技術

多次元範囲検索を可能とするキーバリューストア [MD-HBase]
高倍率・高精細を実現する事例ベースの学習型超解像方式
ビッグデータ活用のためのテキスト分析技術
ビッグデータ時代の最先端データマイニング
ジオタグ付きデータをクラウドでスケラブルに処理するジオフェンシングシステム
柔軟性と高性能を備えたビッグデータ・ストリーム分析プラットフォーム [Blockmon] とその使用事例

◇ 普通論文

地デジ TV を活用した「まちづくりコミュニティ形成支援システム」

◇ NEC Information

NEWS

スケールアウト型新データベース [InfoFrame Relational Store] が 2 つの賞を受賞



Vol.65 No.2
(2012年9月)

特集TOP