

パーソナルデータを暗号化した状態で 安全に匿名化する技術を開発

匿名加工情報の利活用拡大に向けて

さまざまな分野でビッグデータの利活用が進むなか、個人情報保護法の改正案が成立し、将来的にはセンサーやデバイスなどのデータだけでなく、匿名加工情報（個人が特定できないようにしたパーソナルデータ）も第三者が利用できるようになります。これにより、移動履歴や購買履歴といった匿名加工情報を高精度なマーケティングに利用するなど、今後その利活用が大きく拡大すると見込まれています。

こうしたビッグデータの分析ではクラウドの活用が広がって

いますが、パーソナルデータなど機微なデータの取り扱いには、より高い安全性が求められるため、クラウド上のデータを暗号化した状態のまま検索や分析を行う技術の実用化が進んでいます。パーソナルデータを匿名化するk-匿名化技術※1に対しても、セキュリティの向上を目的に、データを暗号化した状態のまま匿名化を行う技術が検討されています。

※1 特定の個人の識別を困難にするためのデータ加工方法の一つ

暗号化状態を維持しながらk-匿名化を実現

そこで日立は、クラウド上のパーソナルデータのセキュリティ向上を目的に、クラウド上では暗号化状態を維持しながらk-匿名化を実現する技術を開発しました。

■安全に暗号化データを集約する技術の開発

多くのk-匿名化技術では、値が異なるデータを同じデータに集約するツリー構造※2を用い、下の階層から上の階層へデータを集約します。しかし、データを暗号化した場合、集約前の下の階層データが暗号化によって解読不能なため、このツリー構造が作成できませんでした。日立は、暗号化状態を維持したままデータの比較を行う独自の検索可能暗号技術を用い、同じ値だと判定された暗号化データの数を集計することでツリー構造を生成する技術を開発しました。

速度が低下するとされています。しかし、日立の検索可能暗号技術を用いると、暗号化データどうしの比較処理を高速に行い、暗号化状態のデータ処理を必要最小限に留めることが可能です。これにより、データ処理のオーバーヘッド増加を30%※3に抑え、実用的な処理速度を確保することに成功しました。

また、より高い安全性を確保するために、データを暗号化する暗号鍵と、暗号化されたデータを匿名化する暗号鍵は、それぞれ異なるものを用います。このため匿名化前の暗号化データが万一流出しても、安全性を担保することができます。

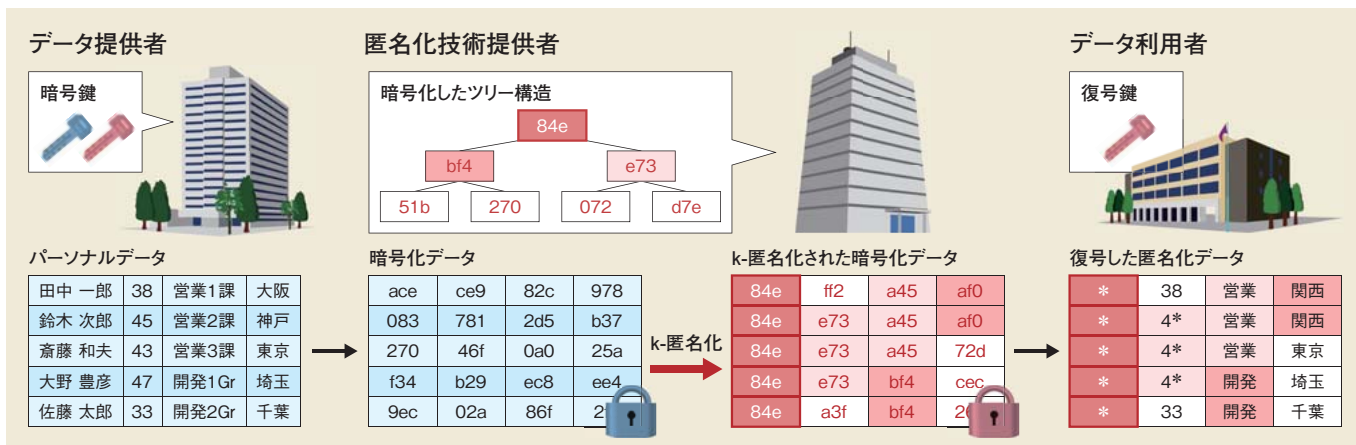
日立は今後、パーソナルデータ利活用の拡大に向けて、2018年度中に今回開発した技術の実用化をめざします。

※2 親となるデータから子となる複数のデータに枝分かれし、さらに孫となる複数のデータに枝分かれするといった形のデータ構造の一種

※3 k値を2に設定し、3万レコード、9種類の属性に対し、暗号化したツリー構造の作成からk-匿名化された暗号化データを作成するまでの、匿名化技術提供者の処理を測定したものです

■実用的な処理速度と高い安全性を達成

一般に、暗号化状態を維持したままのデータ処理は処理



暗号化状態のまま匿名化する技術の概要イメージ

お問い合わせ先

(株)日立製作所 研究開発グループ
<https://www8.hitachi.co.jp/inquiry/hqrd/rd/jp/form.jsp>