(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第4612535号
(P4612535)
(45)発行日　平成23年1月12日(2011.1.12)　(24)登録日　平成22年10月22日(2010.10.22)

(51)Int.Cl.　F I
G06F 13/00 (2006.01)　G06F 13/00　540E

請求項の数 7　(全 10 頁)

(21)出願番号　特願2005-349214(P2005-349214)
(22)出願日　平成17年12月2日(2005.12.2)
(65)公開番号　特開2007-156697(P2007-156697A)
(43)公開日　平成19年6月21日(2007.6.21)
審査請求日　平成20年2月12日(2008.2.12)

(73)特許権者　000004226
日本電信電話株式会社
東京都千代田区大手町二丁目３番１号
(74)代理人　100123788
弁理士　宮崎　昭夫
(72)発明者　荒金　陽助
東京都千代田区大手町二丁目３番１号　日本電信電話株式会社内
(72)発明者　柴田　賢介
東京都千代田区大手町二丁目３番１号　日本電信電話株式会社内
(72)発明者　林　徹
東京都千代田区大手町二丁目３番１号　日本電信電話株式会社内

最終頁に続く

(54)【発明の名称】正当サイト検証手法におけるホワイトリスト収集方法および装置

(57)【特許請求の範囲】
【請求項１】
利用者が受信したメールの文章を解析して企業の正当サイトのＵＲＬホワイトリストを作成する、ホワイトリスト収集装置で行なわれるホワイトリスト収集方法であって、
企業名抽出部が前記メールから企業名を抽出し、取得企業名リストを作成するステップと、
ＵＲＬ解析部が前記メールからＵＲＬを抽出するステップと、
リンク企業名解析部が、抽出された各ＵＲＬおよび抽出された各企業名のメール内の位置関係に基づき、各ＵＲＬごとに、前記取得企業名リストに含まれる、抽出された各企業名に得点を合算し、その得点に応じて企業名を抽出するステップと、
Ｗｅｂサイト取得・解析部が、前記取得企業名リストに含まれる企業名による検索を行い、検索結果である当該企業のトップページＵＲＬおよびそこから辿った範囲のＵＲＬを、当該企業の正当サイトのホワイトリストとして抽出するステップと
を有するホワイトリスト収集方法。
【請求項２】
前記企業名を抽出し、取得企業名リストを作成するステップは、企業名およびその企業名の別名が記載されている企業名リストを用いて、メール内に企業名またはその別名がある場合に、その企業名がメール内に存在したと判断して、その企業名および出現位置を前期取得企業名リストに記録する、請求項１に記載のホワイトリスト収集方法。
【請求項３】

前記企業名を抽出し、取得企業名リストを作成するステップは、３文字以上のアルファベットまたはカタカナを企業名として抽出し、その企業名および出現位置を前記取得企業名リストに記録する、請求項１に記載のホワイトリスト収集方法。

【請求項４】

前記のＵＲＬと深い関係にある企業名を抽出するステップは、メールの文章から抽出されたＵＲＬと、前記取得企業名リストを照合し、メール文章内におけるＵＲＬの出現位置および取得企業名の出現位置から、当該ＵＲＬに対する取得企業名の重み付けを行い、リンク企業名リストに登録する、請求項２または３に記載のホワイトリスト収集方法。

【請求項５】

前記当該会社の正当サイトのホワイトリストとして抽出するステップは、前記取得企業名リストの各企業名に対して、インタネット上の検索サイトによる検索を行い、検索結果最上位のＵＲＬを当該企業のトップページであるとし、当該トップページから一定の範囲のリンクに対して、当該企業の正規サイトとしてホワイトリストに登録する、請求項２または３に記載のホワイトリスト収集方法。

【請求項６】

前記ホワイトリストへの登録範囲として、トップページを基準として同一ドメイン内および同一ドメインから１ホップ内のＵＲＬを登録対象とする、請求項５に記載のホワイトリスト収集方法。

【請求項７】

利用者が受信したメールの文章を解析して企業の正当サイトのＵＲＬホワイトリストを作成する、ホワイトリスト収集装置であって、

前記メールから企業名を抽出し、取得企業名リストを作成する企業名抽出部と、

前記メールからＵＲＬを抽出するＵＲＬ解析部と、

抽出された各ＵＲＬおよび抽出された各企業名のメール内の位置関係に基づき、各ＵＲＬごとに、抽出された各企業名に得点を合算し、その得点に応じて企業名を抽出するリンク企業名解析部と、

検索サイトで抽出した企業名による検索を行い、検索結果である当該企業のトップページＵＲＬおよびそこから辿った範囲のＵＲＬを、当該企業の正当サイトのホワイトリストとして抽出するＷｅｂサイト取得・解析部と

を有するホワイトリスト収集装置。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明はネットワークセキュリティに関し、特にフィッシング詐欺対策方法に関する。

【背景技術】

【０００２】

昨今のネットワーク技術の発達に伴い、メールやＷｅｂなどのネットワークツールを利用して、氏名・年齢・住所・電話番号などの個人情報や、クレジットカード番号・ＩＤ・パスワードなどの信用情報を盗むフィッシング詐欺が多発するようになってきた。

【０００３】

フィッシング詐欺は個人情報の確認や修正を要求する内容のメールを始めとすることが多い。同メールは、あたかも有名な銀行やクレジットカード会社が送信したかのように装い、受信者にＷｅｂサイトにおいて個人情報の入力・修正などを要求し、そのＵＲＬが記載されている。当該ＵＲＬは偽のＵＲＬであるが、当該ＵＲＬの画面は正規のサイトに酷似していたり、アドレスバーが詐称されていたりして、利用者にフィッシング詐欺サイトだと気づかれないよう装っている。図４にこの流れを示す。

【０００４】

このようなフィッシング詐欺への対策として様々な技術が提案されている。

【０００５】

ＣｏｒｅＳｔｒｅｅｔ社のＳｐｏｏｆＳｔｉｃｋは、Ｗｅｂブラウザのツールバーに接続しているサイト名を巨大に表示する手法を取ることで利用者の注意を喚起し、ＩＰアドレスだけといった嫌疑サイトが分かり易いようにしている（非特許文献１）。

【０００６】

Ｄｅｅｐｎｅｔ　Ｔｅｃｈｎｏｌｏｇｉｅｓ社のＤｅｅｐｎｅｔ　Ｅｘｐｌｏｒｅｒは、フィッシングサイトとしてブラックリストに載ったサイトを閲覧しようとした場合や、ＩＰアドレスだけなどのイレギュラーなＵＲＬサイトを閲覧しようとした場合には、ポップアップダイアログにて利用者の注意を喚起する手法を採用している（非特許文献２）。

【０００７】

Ｎｅｔｓｃａｐｅ社のＮｅｔｓｃａｐｅ　８．０は、ブラックリストを用いてフィッシング詐欺サイトを排除する手法を採用している（非特許文献３）。

【０００８】

これらの技術においてはブラックリストを用いてフィッシングサイトの疑いのあるサイトへの接続を防ごうとしているが、以下の理由によりブラックリストを用いた方式の危険性が考えられる。

・全てのフィッシングサイトをインタネット上から発見することは非常に困難

・フィッシングサイトがｏｎｌｉｎｅになってから、それを発見し、ブラックリストに登録するまでには一定の期間が必要である。しかし、フィッシングサイトの平均寿命は６日未満であり、あるフィッシングサイトがブラックリストに登録されてからオフラインとなり消滅するまでの「ブラックリスト有効期間」が非常に短い、またはブラックリスト登録がフィッシングサイト消滅後になる可能性が高い。

・ブラックリストに無いサイトは「危険ではない」と判断されるため、上記のような理由でブラックリストに未登録のフィッシングサイトに対しては効果が全くなく、ブラックリストとのマッチングが失敗（リストに存在しない）した場合にＦａｉｌ　Ｏｕｔなシステムとなっている。

【０００９】

そこで、安全と考えられるサイトを登録するホワイトリスト方式が提案されている（非特許文献４、非特許文献５）。

【非特許文献１】http://www.corestreet.com/spoofstick/

【非特許文献２】http://www.deepnetexplorer.com/

【非特許文献３】http://browser.netscape.com/nsb/support/relnotes.jsp

【非特許文献４】SecureBrain 社のPhishWall、http://www.securebrain.co.jp/products/phishwall/index.html

【非特許文献５】NetMove 社のnProtect Netizen、http://nprotect.jp/netizen/

【発明の開示】

【発明が解決しようとする課題】

【００１０】

しかしながら、これらのホワイトリスト方式では、ホワイトリストを収集する手法が課題となる。上記のＳｅｃｕｒｅＢｒａｉｎ社、ＮｅｔＭｏｖｅ社とも特定の企業と契約することで、その企業のサイトをホワイトリストに登録する手法を採っている。しかし、エンドユーザの視点では、小数の契約企業のサイトだけではなく、より多くのサイトに対してホワイトリストを発行することが望まれる。

【００１１】

また、従来の手法では、契約企業のサイト構成を把握し、ユーザ端末に対してそれを証明する役割のセンタが必要であり、一定の運用コストが必須である。

【００１２】

本発明の目的は、ホワイトリスト管理・証明用のセンタが不用なホワイトリスト収集方法および装置を提供することにある。

【課題を解決するための手段】

【００１３】

上記目的を達成するために、本発明のホワイトリスト収集方法は、
利用者が受信したメールの文章を解析して企業の正当サイトのＵＲＬホワイトリストを作成する、ホワイトリスト収集装置で行なわれるホワイトリスト収集方法であって、
企業名抽出部が前記メールから企業名を抽出し、取得企業名リストを作成するステップと、
ＵＲＬ解析部が前記メールからＵＲＬを抽出するステップと、
リンク企業名解析部が、抽出された各ＵＲＬおよびそれらのメール内の位置関係から、抽出された企業名のうち当該ＵＲＬと関係の深い企業名を抽出するステップと、
Ｗｅｂサイト取得・解析部が、取得企業名リストに含まれる企業名による検索を行い、検索結果である当該企業のトップページＵＲＬおよびそこから辿った範囲のＵＲＬを、当該企業の正当サイトのホワイトリストとして抽出するステップと
を有する。

【００１４】

本発明では、フィッシング詐欺の発端となるフィッシングメールの内容を解析して詐称先企業名を割り出し、インタネット上の検索エンジンを利用して当該企業名から当該企業名の正規サイトを発見し、そのサイトをクロールして情報を集めることで、ホワイトリストをユーザ端末が自動的に取得するものである。

【発明の効果】

【００１５】

ユーザ端末が受信したメールの情報と検索サイトを用いてホワイトリストを構築するため、従来のホワイトリストでは必須であったホワイトリスト管理・証明用のセンタが不要となり、運用コスト削減に資する。

【発明を実施するための最良の形態】

【００１６】

次に、本発明の実施の形態について図面を参照して説明する。

【００１７】

図１は本発明の一実施形態によるホワイトリスト収集装置のブロック図、図２は図１のホワイトリスト収集装置の処理の流れを示すフローチャートである。

【００１８】

本実施形態のホワイトリスト収集装置はメール解析部１とホワイトリスト構築部２で構成されている。

【００１９】

メール解析部１は、利用者が受信したメール３を解析し、メール３中に現れる企業名およびメール中のＵＲＬと結びつく企業名を解析・抽出し、取得企業名リスト１４およびリンク企業名リスト１８を作成するもので、ＵＲＬルール辞書１１とＵＲＬ解析部１２と企業名ルール辞書１３と企業名リスト１４と企業名解析部１５と取得企業名リスト１６とリンク企業名解析部１７とリンク企業名リスト１８で構成されている。

【００２０】

企業名解析部１５は、企業名ルール辞書１３および企業名リスト１４を用いてメールヘッダとメールボディ（ヘッダ以外）から企業名を抽出する（ステップ１０１）。企業名ルール辞書１３は、例えば正規表現で記述されており、表１、表２に示すように、「変数名リスト」と「ルールリスト」から構成される。

【００２１】

【表１】

変数番号	変数名	変数値
v1	$ZA	'(?:¥xA3[¥xC1-¥xDA¥xE1-¥xFA])'
v2	$HA	'(?:[a-zA-Z])'
v3	$ZK	'(?:¥xA5[¥xA1-¥xF6]¥xA1[¥xA6¥xBC¥xB3¥xB4])'

【００２２】

【表２】

ルール番号	ルール	企業名変数
r1	@list = ($_ =~ m/($ZA{3,})/g);	@list
r2	@list = ($_ =~ m/($HA{3,})/g);	@list
r3	@list = ($_ =~ m/($ZK{3,})/g);	@list

【００２３】

企業名解析部１５は「変数名リスト」に記載されている“変数値”を“変数名”で記憶する。そして「ルールリスト」に記載されている各“ルール”の変数を、記憶した“変数名”－“変数値”の組で展開する。そして、入力である、メールヘッダとメールボディに対して“ルール”を用いて企業名を抽出する。なお、表２に示す例では、正規表現解析対象である入力は“＄＿”に代入されているとしている。そして、“企業名変数”に格納される抽出された企業名を取得企業名リスト１６に登録する（ステップ１０２）。例として表１に示す変数ｖ１、ｖ２、ｖ３は、それぞれＥＵＣ　ＪＰ　Ｅｎｃｏｄｉｎｇの「全角アルファベット」「半角アルファベット」「全角カタカナ」を示し、表２に示すルールｒ１、ｒ２、ｒ３は、それぞれ「３文字以上の連続する全角アルファベット」「３文字以上の連続する半角アルファベット」「３文字以上の連続する全角カタカナ」を企業名として抽出するルール例である。

【００２４】

表３に企業名リスト１４の表記例を示す。

【００２５】

【表３】

項番	企業名	別名
1	ABC株式会社	ABC,ABC Corp.,ABC(株), ABC株式会社
2	XYZ銀行	XYZ Bank,XYZ銀行,XYZ
3	あいう生命保険相互会社	あいう生命,あいう生命保険,あいう生命保険相互会社

【００２６】

“企業名”およびその“別名（エイリアス）”から構築されており、“企業名”または“別名”に合致する文字列をメールヘッダとメールボディに見つけた場合には、それを取得企業名リスト１６に登録する。なお、“別名”で抽出された場合には、その“別名”の“企業名”で取得企業名リスト１６に登録されることで、複数の書き方が乱立した場合でも統合することが可能となる。さらに、企業名ルール辞書１３を用いて抽出された企業名についても企業名リスト１４との照合を行い、それが“別名”として登録されている場合には、その“別名”の“企業名”で取得企業名リスト１６に登録が行われる。取得企業名リスト１６の表記例を表４に示す。

【００２７】
【表４】

項番	抽出企業名	出現位置	
		行	列
1	ABC株式会社	3	10
2	ABC株式会社	10	2
3	XYZ銀行	13	6
4	ABC株式会社	18	20
5	あいう生命保険相互会社	23	31

【００２８】
企業名解析部１５により抽出された企業名とその企業名が現れる位置（行数とその行の何文字目か）が記憶されている。
【００２９】
一方、ＵＲＬ解析部１２は、メールヘッダとメールボディの中からリンクとなるＵＲＬと思われる文字列を抽出する（ステップ１０３）。表５にＵＲＬ解析部１２が利用するＵＲＬルール辞書１１の記述例を示す。図３の例では、「ｈｔｔｐ：／／」または「ｈｔｔｐｓ：／／」または「ｆｔｐ：／／」で始まるアドレスをＵＲＬであると定義している。
【００３０】
【表５】

項番	URL	企業名リスト	
		企業名	得点
1	http://abc_corp.co.jp	ABC株式会社	30
		XYZ銀行	12
		あいう生命保険相互会社	9
2	http://www.xyz-bank.co.jp/	XYZ銀行	26
		あいう生命相互保険会社	12

【００３１】
ＵＲＬ解析部１２は抽出したＵＲＬをリンク企業名解析部１７に通知する。
【００３２】
リンク企業名解析部１７は、取得企業名リスト１６とＵＲＬ解析部１２から通知されるＵＲＬリストおよびそれらの位置を用いて、各ＵＲＬとそれが指し示す企業名との関係を求める（ステップ１０４）。例えば、メールヘッダのＳｕｂｊｅｃｔに記載されている企業名は＋２０点、当該ＵＲＬ前５行以内に出現する企業名は＋１０点、メール末尾に記載されている企業名は＋１５点などのルールに基づき得点が合算され、各ＵＲＬに対して関係があると考えられる企業名のリストとその得点を抽出し、リンク企業名リスト１８に格納される（ステップ１０５）。リンク企業名リスト１８の記述例を表６に示す。
【００３３】

【表６】

項番	ルール
1	http://[¥d¥w¥.¥_¥/¥~]+
2	https://[¥d¥w¥.¥_¥/¥~]+
3	ftp://[¥d¥w¥.¥_¥/¥~]+

【００３４】

ホワイトリスト構築部２はホワイトリストを構築するもので、Ｗｅｂサイト取得・解析部２１とホワイトリスト２２から構成される。

【００３５】

Ｗｅｂサイト取得・解析部２１は取得企業名リスト１６にある企業名に対応するサイトを検索・収集してホワイトリスト２２を作成する。Ｗｅｂサイト取得・解析部２１は、取得企業名リスト１６の“抽出企業名”を用いて、インタネット上の検索サイトで検索処理を行う（ステップ１０６）。そして、検索結果のトップを当該企業のホームページであるとして取得し（ステップ１０７）、ホワイトリスト２２にトップページとして登録する（ステップ１０８）。次に、トップページのソースを解析し、トップページからリンクが張られているページを取得する。そして、さらにそのページのリンクについても同様に回帰的にページを取得してゆく。ただし、トップページのＵＲＬを解析し、そのＵＲＬと関係のあるページまでしか取得しないこととする。“関係のある”とは、例えば以下のような定義に基づく。

・トップページと同じドメイン内であれば関係があるとする。

・トップページと同じドメインであっても、トップページが「ｈｔｔｐ：／／ｗｗｗ．ｉｓｐ．ｎｅ．ｊｐ／～ｕｓｅｒｎａｍｅ／」などの“～（チルダ）”を利用している場合には、チルダの示すアカウント名のディレクトリ配下は関係があるとする。

【００３６】

ホワイトリスト登録処理例を図３に示す。他ドメインであっても同一ドメインからリンクが１ホップにあるＵＲＬは信頼して「ホワイトリスト」に登録するが、２ホップ以上についてはホワイトリストには登録しない。ただし、図３に示すようにリンクを辿る経路において再び同一ドメインとなった場合は「ホワイトリスト」に登録するものとする。

【００３７】

「ホワイトリスト」の記述例を表７、表８に示す。

【００３８】

【表７】

項番	企業名	トップページ
top01	ABC株式会社	http://www.abc_corp.co.jp/
top02	XYZ銀行	http://www.xyz-bank.co.jp/

【００３９】

【表８】

項番	トップページ	ホワイトリスト
wl01	http://www.abc_corp.co.jp/	http://www.abc_corp.co.jp/gnews/index.html http://research.abc_corp.co.jp/jump/caution.html http://www.abc_corp.co.jp/RD/OFIS/top.html
wl02	http://www.xyz-bank.co.jp	http://www.xyz-bank.co.jp/retail/ http://www.xyz-bank.co.jp/company/

【図面の簡単な説明】

【００４０】

【図１】本発明の一実施形態によるホワイトリスト収集装置のブロック図である。

【図２】図１のホワイトリスト収集装置の処理の流れを示すフローチャートである。

【図３】ホワイトリストに登録する範囲を示す図である。

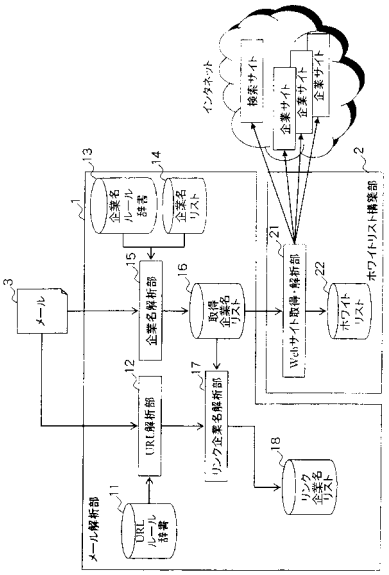
【図４】フィッシング詐欺の流れの説明図である。

【符号の説明】

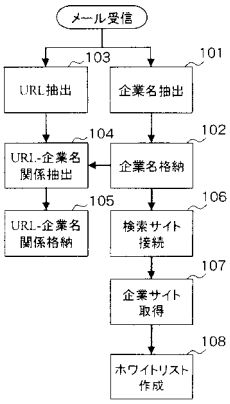
【００４１】

１　　メール解析部
２　　ホワイトリスト構築部
１１　　ＵＲＬルール辞書
１２　　ＵＲＬ解析部
１３　　企業名ルール辞書
１４　　企業名リスト
１５　　企業名解析部
１６　　取得企業名リスト
１７　　リンク企業名解析部
１８　　リンク企業名リスト
２１　　Ｗｅｂサイト取得・解析部
２２　　ホワイトリスト
１０１～１０８　　ステップ

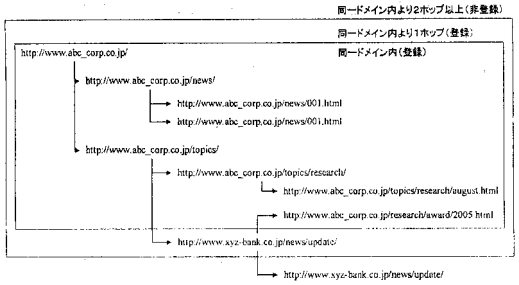
【図１】


インターネット
検索サイト
企業サイト
企業サイト
企業サイト
2
ホワイトリスト構築部
13
企業名ルール辞書
14
企業名リスト
1
21
3
メール
15
企業名解析部
16
取得企業名リスト
Webサイト取得・格納部
22
ホワイトリスト
12
URL解析部
17
リンク企業名解析部
18
リンク企業名リスト
メール解析部
11
URLルール辞書

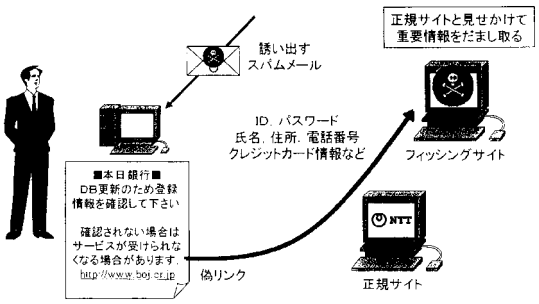
【図２】


メール受信
103
101
URL抽出
企業名抽出
104
102
URL-企業名関係抽出
企業名格納
105
106
URL-企業名関係格納
検索サイト接続
107
企業サイト取得
108
ホワイトリスト作成

【図３】


同一ドメイン内より2ホップ以上（非登録）
同一ドメイン内より1ホップ（登録）
同一ドメイン内（登録）
http://www.abc_corp.co.jp/
http://www.abc_corp.co.jp/news/
http://www.abc_corp.co.jp/news/001.html
http://www.abc_corp.co.jp/news/001.html
http://www.abc_corp.co.jp/topics/
http://www.abc_corp.co.jp/topics/research/
http://www.abc_corp.co.jp/topics/research/august.html
http://www.abc_corp.co.jp/research/award/2005.html
http://www.xyz-bank.co.jp/news/update/
http://www.xyz-bank.co.jp/news/update/

【図４】


誘い出す
スパムメール
正規サイトと見せかけて
重要情報をだまし取る
ID、パスワード
氏名、住所、電話番号
クレジットカード情報など
フィッシングサイト
■本日銀行■
DB更新のため登録
情報を確認して下さい
確認されない場合は
サービスが受けられな
くなる場合があります。
http://www.boj.or.jp
偽リンク
NTT
正規サイト

フロントページの続き

審査官　木村　雅也

(56)参考文献　特開２００５－１３５０２４（ＪＰ，Ａ）
国際公開第２００５／１０９２２５（WO，Ａ１）
国際公開第２００３／０４６７６４（WO，Ａ１）
柴田　賢介，フィッシング詐欺対策のためのＵＲＬ検証方式の提案，マルチメディア，分散，協調とモバイル（ＤＩCOMO）シンポジウム論文集　１９９７年～２００６年版　Ｖｅｒ．１．１　［ＤＶＤ－ＲOM］，日本，社団法人情報処理学会，２００５年　7月　6日，第2005巻，第485頁-第488頁
荒金　陽助，フィッシング詐欺対策に向けた一考察，マルチメディア，分散，協調とモバイル（ＤＩCOMO）シンポジウム論文集　１９９７年～２００６年版　Ｖｅｒ．１．１　［ＤＶＤ－ＲOM］，日本，社団法人情報処理学会，２００５年　7月　6日，第2005巻，第481頁-第484頁

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ　１３／００
Ｇ０６Ｆ　１５／００