Um estudo sobre métodos de pesquisa utilizados em segurança computacional - Criptografia

Introdução

O atual modelo de produção científica tem como maior prioridade a promoção dessa produção através de artigos publicados em eventos científicos, o qual foram referendados por sua qualidade através de um corpo editorial.

Não há uma forma explícita de escrever a forma como foi produzido o conhecimento do trabalho de pesquisa. Contudo, um modelo convencional para artigos científicos na Ciência da Computação inclui, tradicionalmente, os seguintes elementos [23]:

1. trabalho descreve uma nova idéia, talvez prototipada em um pequeno sistema; e/ou
2. trabalho alega o seu lugar na Ciência, realizando comparações de características - ou seja, o relatório apresenta uma lista de características, e compara qualitativamente antigas abordagens com a nova, característica a característica.

Tal artigo poderia ser considerado satisfatório, se vem a apresentar uma idéia realmente radical, ou uma quebra de paradigma. Talvez, quando pela primeira vez se realizou um experimento usando Java, ou ao se propor um Navegador Web, tenha havido essa mudança ou descoberta.

Não obstante, não é o que se observa ao estudar os tipos de métodos de pesquisa. Encontramos, muitas vezes, artigos sem fundamentação teórica, sem a definição de métodos de pesquisa, e diversos outros problemas. Se o corpo editorial dos anais de eventos onde cada artigo publicado verificasse e metodologia empregada, poder-se-ia obter trabalhos com maior rigor e maior qualidade científica. Hoppen [8] destaca a relevância de destacar o método empregado em pesquisas, mencionando-os nos artigos publicados.

O presente trabalho apresenta o resultado de uma pesquisa *survey* realizada em fins do segundo semestre de 2000, onde os objetos de pesquisa foram artigos publicados em eventos internacionais de Criptografia, que dispunham de corpo editorial. O objetivo de tal estudo foi identificar os tipos de métodos de pesquisa empregados nessa área, além de verificar a existência ou não de correlações geográficas ou temporais, com relação ao método empregado.

Esse estudo encontra-se organizado da seguinte forma: a seção 2 apresenta a metodologia geral da pesquisa, deixando para a seção 3 a descrição de diversos métodos de pesquisa, tradicionalmente empregados nas Ciências. A seção 4 apresenta o método empregado no presente trabalho, bem como os resultados das análises dos dados obtidos. Na última seção, algumas considerações finais e recomendações para trabalhos futuros são apresentadas.

Métodos de Pesquisa

A presente seção apresenta a metodologia geral da pesquisa, os tipos de métodos tradicionalmente aplicados em Ciências, e os tipos de métodos que recentemente têm tido mais progressos, graças ao advento dos computadores, de modo genérico. Aspectos particulares de cada um dos métodos mais utilizados – os chamados tradicionais - serão apresentados em capítulo posterior.

Metodologia Geral da Pesquisa

O trabalho científico inicia com uma pergunta a ser respondida, um problema, uma questão de pesquisa. Há uma dúvida, uma pergunta, uma questão, um problema a ser resolvido [KER 79].

Segue-se a necessidade de haver uma idéia que a justifique. Talvez, um conceito a ser definido. Em muitos casos, usa-se ou cria-se um construto [8]. Para garantir a isenção e buscar o veracidade dentro do possível, usa-se de toda uma formalização da pesquisa – pois parte-se do princípio (de modo geral) que outros investigadores, dadas as mesmas condições, poderiam repetir tais formalismos, buscando obter resultados semelhantes. Para a formalização de uma pesquisa, faz-se uso de métodos. O método a ser utilizado depende do tipo de pesquisa, conforme será visto adiante.

Diversos autores buscam formalizar o próprio procedimento de pesquisar, decompondo o processo em fases, e interligando essas fases. Essa ação facilita o processo de ilustrar graficamente todo o processo de pesquisa, possibilitando uma visualização mais adequada de todo o processo.

Na realidade, o estilo pode a vir influenciar todo o processo de condução de pesquisa; uma intervenção não terá todo o formalismo de uma pesquisa por enquetes – em contrapartida, deverá deixar claro e evidente em seus relato a forma como foram obtidos ou selecionados os indivíduos, como se coletaram os dados, e principalmente, como se chegou às conclusões apresentadas.

Tipos de estudos científicos

Os estudos científicos têm características específicas no que tange ao seu caráter ou tipo. De forma geral, os estudos podem ter caráter exploratório, descritivo, explanatório ou preditivo. Exploratório quer dizer que há novos aspectos, novos enfoques ou até novas áreas a serem descobertas, podendo abrir campo para muitas e novas pesquisas. Descritivo significa que já foram feitos estudos exploratórios, permanecendo dúvidas com relação a certos fenômenos ou sujeitos de pesquisa, os quais devem ser descritos. Explanatório significa descrever relações causa/efeito. Preditivo pretende colocar possíveis cenários ou conseqüências.

Os trabalhos científicos podem ter aplicação imediata, ou de forma contrária, nem mesmo se vislumbrar alguma aplicação para o fenômeno observado ou para o resultado da pesquisa.

Pode-se observar que o tipo de pesquisa depende do enfoque ou critério pretendido pelo autor, podendo obedecer a interesses em particular (ou a determinadas condições). Pode-se classificar os tipos de pesquisas de diversos modos, conforme o quadro a seguir:

Contudo, observa-se ainda não haver unanimidade em definir os critérios. Ciências de caráter social enfatizam alguns tipos, ao passo que nas Exatas notamos alguns comuns às Ciências anteriores – havendo contudo outros específicos.
A seguir, realiza-se breve descrição dos tipos:
1 – Básica – procura ampliar o campo de conhecimento teórico – sem, no entanto preocupar-se com a imediata aplicabilidade dos resultados. Seus resultados podem ser leis, princípios ou generalizações.
2 – Aplicada – busca resultados que possam ser utilizados na solução de problemas reais [13].
3 – Descritiva – busca descrever situações, eventos ou fenômenos. O estudo descritivo tem como requisito o conhecimento prévio considerável da área de estudo, para que se possa formular as perguntas de pesquisa [19][13].
4 – Exploratório – procura examinar um tema ou área pouco estudado, ou sob um novo enfoque. Sua utilização é aumentar o grau de familiaridade com fenômenos ou eventos pouco conhecidos [19] [MAR90].
5 – Correlacional – busca estabelecer relações entre dois ou mais conceitos, ou o grau de relação entre esses conceitos. O principal objetivo é determinar como se comporta um conceito, na presença de outro. Bastante similar ao tipo "Explanatória", onde se busca estabelecer a relação causa-efeito enter os conceitos [19].
6 – Preditivos – procura predizer os resultados de um fenômeno, ou ainda, seus eventos e comportamentos.
De modo geral, os estudos podem iniciar com um tipo, mas não se situar em apenas um dos tipos. O pesquisador pode descobrir outros enfoques durante o trabalho de pesquisa. Pode ocorrer que pesquisadores pretendam conduzir uma pesquisa dentro de determinado enfoque e seja obrigado a direcionar para outro - por exemplo, um estudo inicialmente exploratório pode ser obrigado a redirecionar os esforços para um estudo descritivo, em vista de haver descoberto durante o processo de pesquisa que seu trabalho já teve precedentes.

Métodos de Pesquisa

Classicamente, autores [19] [8] [13]dividem os métodos de pesquisa em sete tipos de métodos de pesquisa: a bibliográfica, o *survey*, a experimentação, o estudo de caso, a histórica, a pesquisa-ação.
Os quatro primeiros incorrem em uma postura menos participativa do pesquisador. Assim sendo, parte-se sempre do princípio de que há uma nítida separação entre o objeto a ser estudado e o pesquisador. Não é permitida nenhuma forma de interação entre os dois: isso incorreria em uma forma de parcialidade ou influência nos resultados da pesquisa, o que inviabilizaria o trabalho de pesquisa. Tem ampla influência positivista.
O quinto pode, por vezes, admitir a participação do pesquisador – embora não o seja possível no contexto temporal. Destaca-se que nem sempre o pesquisador, nesse método, assume uma postura distante do objeto – embora utilize, muitas vezes, artifícios que os métodos com influência positivista empregam. É, muitas vezes, dito seguir a abordagem do paradigma interpretativo.
O sexto pressupõe ampla participação – e até convivência – do pesquisador no problema. Parte-se do princípio que o pesquisador faz parte do contexto do problema de pesquisa, e assim sendo, deve ser citado no trabalho, com todas as informações pertinentes as suas intervenções. Por esse motivo, essa metodologia também é conhecida por intervenção. Diferente das outras metodologias, não se parte de hipóteses, mas de pressupostos. É uma metodologia muito utilizada em pesquisa social – em especial, na Antropologia. Diz-se fazer parte do paradigma crítico.
A Ciência da Computação, por se tratar de uma Ciência muito recente, vai adquirindo ou agregando metodologias amplamente aplicadas em outras Ciências – razão pela qual podemos observar diversas metodologias não clássicas, como será descrito em seção posterior.

Métodos de Pesquisa

A seção a seguir apresenta – na primeira parte -,os principais métodos de pesquisa empregados nas Ciências Exatas de forma tradicional. São eles o estudo de caso, a pesquisa *survey*, e a experimentação. Para o caso de haver alguma diferença notável em alguma fase da pesquisa – como, por exemplo, uma forma diferente de realizar a validação -, será indicada no desenvolvimento do item. Posteriormente, são apresentados os métodos de pesquisa que vêm se desenvolvendo graças ao advento dos recursos computacionais.

O Estudo de Caso

Trata-se, sem dúvida, de um dos métodos utilizados em Ciência da Computação – e em outras Ciências, como Sociais, Exatas, Médicas, etc. A decisão de se usar o método mais apropriado depende do que será investigado. Esse "do que será investigado" trata-se da questão de pesquisa, a qual será definida sempre antes de qualquer outro trabalho prático.
Pode ser considerada como sendo uma ***estratégia*** de pesquisa [24].
O foco nesse método é estudar PROFUNDAMENTE UMA unidade, no seu ambiente natural, a partir de diversas fontes de evidência, pelo emprego de diversas técnicas de coleta de dados [8] [24]. Para tanto, é fundamental que:
A - O pesquisador não manipule os indivíduos pertencentes às fontes de evidência; e
B – O pesquisador não exerça controle nenhum sobre os indivíduos.
Normalmente, os fenômenos que possam ser estudados podem estar no curso de sua história, ou podem ser estudados sobre o seu passado.
Há algumas situações onde o estudo de caso é a única forma de levar a cabo um estudo. Nem sempre podemos ter uma quantidade de indivíduos que justifique uma pesquisa por enquetes – e, se os tivermos, e tivermo-los à disposição do pesquisador, nem sempre nos interessam os dados da coletividade. No caso de uma experimentação, não podemos colocá-los sob o total controle. Pode nos interessar os fatos mais recente – o que exclui a pesquisa histórica. Pode não haver fontes bibliográficas que nos resolvam o problema – pesquisa puramente bibliográfica, descartada.

Afora problemas com o método, podemos ter problemas com indivíduos (por exemplo, o caso do mamífero australiano chamado tilacine – só havia um exemplar vivo, em 1936), os quais podem ser únicos.
Finalmente, podemos apenas ter a profundidade necessária no caso de presenciar o fenômeno junto ao indivíduo [24].
Há 3 condições básicas[24] para executar um estudo de caso:

- tipo de questão de pesquisa colocada;
- grau de controle que um investigador tem sobre os eventos; e
- quão contemporâneo é o estudo
- estudo de caso é uma necessidade empírica quando

A - investiga um fenômeno contemporâneo no contexto de sua vida real, especialmente quando
B - os limites entre o fenômeno e o contexto não estão claramente evidentes.
A condução de um estudo de caso engloba diversas fases: inicia com a definição do problema de pesquisa – do qual deriva a pergunta de pesquisa - ; passa-se, então, para a elaboração de construtos. Pode-se efetuar uma pesquisa bibliográfica, a fim de melhor definir o referencial teórico. Decide-se, então quais serão os elementos a pesquisar – quem são os indivíduos, e/ou a unidade de pesquisa. Projeta-se a condução do Estudo de Caso. Passa-se à coleta de dados, e sua interpretação e análise [4]. Daí, para as conclusões.
A questão de pesquisa normalmente é formulada como premissa que orienta o estudo, tendo em vista que uma formulação de hipóteses é muito restrita para guiar as ações e não se aplica a pesquisas qualitativas, principalmente quando a posição epistemológica for interpretativa. Na análise de estudos que utilizam metodologias qualitativas, a avaliação das bases teóricas é fundamental. Quando a abordagem for positivista e o objetivo do estudo for a verificação de uma teoria, muitas vezes é necessário enunciar as teorias rivais existentes sobre o fenômeno em análise, para permitir uma real comprovação[8]. A questão de pesquisa normalmente é de um dos tipos a seguir: "quem", "o que/qual/quais", "onde", "como" e "porquê".
De modo geral, questões de pesquisa com principal foco em "o que/qual/quais" podem ter duas possibilidades: "Quais são os modos de tornar uma interface mais amigável?" Esse tipo de questão é uma justificativa razoável para conduzir um estudo *exploratório*, para desenvolver hipóteses pertinentes e proposições para questões futuras. Ademais, encontra-se embutido ao menos um problema a ser solucionado. Porém, essa questão também poderia ser respondida por outros métodos. O outro tipo de questão é, na verdade, um modo de linha de questionamento do tipo "quantos", ou "quanto". "Quais têm sido o tipo de resultados do impacto do uso de um banco de dados relacional na equipe de desenvolvimento? " Normalmente, tal tipo de pergunta beneficia – em uma primeira vista – outros tipos de estratégias, como um *survey* (que pode enumerar "quais"). Similarmente, como esse segundo tipo de pergunta "o que/qual/quais", perguntas tipo "quem " ou " onde" são desejáveis em pesquisas por enquetes, ou até em análise de registros históricos, como em uma pesquisa que desejar argumentar aspectos micro-econômicos ou macro-econômicos. Essas estratégias são vantajosas quando o objetivo da pesquisa é descrever a incidência ou o predomínio de um fenômeno, ou ainda, quando se pretende predizer determinadas possibilidades de resultados. Um levantamento sobre atitudes que um programador tem, ao ser obrigado a adotar outro paradigma ou outra linguagem de programação (onde um *survey* ou um censo pode ser a estratégia mais conveniente), ou o decréscimo de uso de "sistemas de informação" na qual se usou Clipper ou COBOL(na qual a análise de dados históricos é a estratégia mais conveniente) podem ser exemplos típicos. Em contraste, questões do tipo "como" e porque" são mais *explanatórias*, e desejáveis de conduzir um estudo de caso – embora também possam ser utilizadas em estudos experimentais. Isso ocorre em função de que as questões levam a "elos" (links) operacionais com outras idéias ou construtos.
Chama-se atenção para o fato de que algumas perguntas do tipo "como" e "por que" são ambivalentes, e podem necessitar clarificação. Muita atenção é requerida para a formulação da pergunta de pesquisa.
Embora possa-se e deva-se usar diversas fontes de busca de evidências para a composição do estudo de caso, as mais freqüentemente usadas nos estudos de caso são a observação direta e entrevistas sistemáticas.
De qualquer modo, ressalta-se que se pode usar outras fontes, como registros históricos, fotografias (em especial, nas pesquisas da Geociências), artefatos (principalmente, nos estudos realizados nas Engenharias), levantamentos etc. Destaca-se a necessidade de SEMPRE se descrever a forma como foi operacionalizada a busca de evidências e a coleta de dados. Observa-se que, em alguns casos, a manipulação informal pode ocorrer, ao se usar a técnica de observação participante.
Uma investigação tipo estudo de caso

- parte da situação distinta (tecnicamente) na qual haverá muito mais variáveis de interesse do que pontos de dados, chegando a um resultado;
- baseia-se em múltiplas formas de evidência, com dados necessitando convergir em uma forma triangular, pontual ou não convergente;
- beneficia a partir do primeiro desenvolvimento de uma preposição teórica, para guiar a coleta de dados e posterior análise dos mesmos.

Assim como no caso da pesquisa-ação (ou intervenção), na observação participante, e de alguns casos da dissertação-projeto, as técnicas aplicadas na metodologia variam muito. Muitas vezes, para se beneficiar, o pesquisador poderá adotar mais de uma metodologia – porém, sempre mantendo a mesma postura. Qual a vantagem disso? Assim, poderá

ele reunir o maior número possível de pontos de vista, bem como outras informações pertinentes, para melhor observar o objeto de estudo. A essa abordagem, chamamos de multi-métodos (conforme Norberto Hoppen [HOP96]).
Os componentes de um projeto (ou desenho) de pesquisa são:

- questão/pergunta de pesquisa;
- proposições – se existirem;
- unidades de análise;
- ligação lógica dos dados com as proposições; e
- critérios para interpretar as afirmações.

Destaca-se a preocupação com o referencial teórico: o desenvolvimento de teorias fundamentadas e validadas corretamente tem sido raro. O desenho a seguir apresenta o caso da generalização – tão complexo no estudo de caso -, tendo por base estudos estatísticos, baseados em *survey* ou em experimentações.
Após obtidos os dados, o pesquisador normalmente empreende 3 fases: a codificação dos dados, a apresentação de modo a atender ao projeto, e a análise dos mesmos.

A Pesquisa Survey

A pesquisa denominada "*survey*" é também conhecida como pesquisa por enquetes, ou ainda pesquisa por levantamentos. Parte-se do princípio de obter o conhecimento através da quantidade. Assim, é um método quantitativo, e tem seu apoio na Estatística. Freqüentemente, busca levantar um perfil, e é indicado para identificar acontecimentos recentes ou em um passado próximo, ou em pesquisas de opinião.
O item a seguir apresenta a pesquisa *survey*, os objetivos, cuidados a serem empregados, o processo de amostragem, a coleta e análise de dados.
A primeira atividade a realizar – em qualquer forma ou método de pesquisa, é identificar o problema de pesquisa Uma vez identificado, pode-se identificar pressupostos relativos ao problema. A seguir, busca-se esclarecer a(s) justificativa(s) ou relevância. Uma das formas mais simples de estabelecer a relevância é usar dados bibliométricos.
Dentre uma série de temas a estudar, pode-se alegar o aumento de interesse desse tema em algum intervalo de tempo como uma forma de justificar a relevância do tema.
O próximo passo é buscar estabelecer os objetivos da pesquisa. Devem ser definidos tanto o objetivo principal ou geral do estudo, quanto os específicos. A definição de objetivos específicos tanto pode colaborara para compor o objetivo geral, quanto para abrir novas frentes de pesquisas ou trabalhos futuros. O método de pesquisa *survey* é bastante difundido em algumas áreas da Computação, - tal como a área de sistemas de informação- , de acordo com Orlikowski e Baroudi [[15]. A mesma informação é colocada por VOGEL e WETHERBE (*apud* Pinsonneault e Kraemer [17]). Já Pinsonneault e Kraemer [17] caracterizam a natureza da pesquisa *survey* comparando-a com os outros dois métodos dominantes na pesquisa na área de sistemas de informação: o estudo de caso e os experimentos de laboratório. Contrastando com os dois métodos, os autores lembram que a pesquisa *survey* envolve o exame de um fenômeno em uma grande variedade de questões. O pesquisador claramente define as variáveis dependentes e independentes, assim como um modelo específico das relações esperadas, as quais são testadas com as observações do fenômeno (Pinsonneault e Kraemer [17]).
Ainda segundo os últimos, a pesquisa *survey* é mais apropriada quando:

1. as questões centrais de interesse sobre o fenômeno são "o que está acontecendo", ou "como e por que está acontecendo" - ou seja, é indicada para responder a perguntas sobre "o quê", "com" e "quantos"-;
2. não é desejável ou possível controlar as variáveis - sejam dependentes ou independentes -;
3. fenômeno deve ser estudado em seu ambiente natural;
4. fenômeno ocorreu num passado recente ou está ocorrendo.

Para Marconi e Lakatos [MAR 92], o desenho de pesquisa é uma forma esquemática que facilita a viabilidade de uma pesquisa, podendo ou não ser modificado. O desenho auxilia o pesquisador a imprimir uma ordem lógica no trabalho. Muitas vezes, identifica plenamente o construto, e o seu uso.
O problema de pesquisa deve ser solucionável em determinado prazo. Assim, o pesquisador deve estabelecer, dentro da realidade que lhe é proporcionada, um determinado tempo para a realização do trabalho de pesquisa. Ao decompor as atividades componentes do trabalho de pesquisa, estimando o tempo necessário para a realização de cada tarefa, e colocando no tempo de realização da pesquisa, obtém-se um cronograma. O cronograma deve ser encarado, antes de tudo, como uma ferramenta para o apoio do pesquisador.
Para a definição das indivíduos que fazem que compõe determinada amostra, deve ser realizado o processo de amostragem. O mesmo é realizado tendo, por início, a definição do universo com que se pretende trabalhar.
O passo seguinte procura identificar o(s) critério(s) de seleção de indivíduos, ou critério(s) de amostragem. Baseado nesses critérios, passa-se então a descrever a composição de toda a amostra.
Por vezes, torna-se interessante destacar se existem, no universo, indivíduos com características distintas. Muitas vezes, essas características podem estabelecer interessantes relações entre as variáveis de pesquisa – seja por agrupamento, seja por estratificação dos grupos de indivíduos com características distintas.
Sampieri [19] afirma que, após selecionar o tipo de pesquisa, e a amostra, o passo seguinte é coletar os dados pertinentes às variáveis envolvidas na investigação. A coleta de dados implica três atividades estreitamente vinculadas entre si:

a. a seleção de um instrumento de medida dos dados - que deverá ser válido e confiável;
b. a aplicação desse instrumento - a obtenção das medidas e observações das variáveis de interesse em nosso estudo; e
c. a preparação dessas medidas para que se possa analisar corretamente.

Em um trabalho do tipo *survey*, o instrumento típico a ser empregado é o questionário - cuja composição do tipo das variáveis é descrito adiante -, e sua estratégia de aplicação pode ser das mais diversas – tendo o correio eletrônico facilidades interessantes . Escalas empregadas e medidas aplicadas serão apresentadas nos próximos itens.

Os erros que podem ocorrer em pesquisas de sistemas de informações são de dois tipos: os erros amostrais e os erros não amostrais. O erro total de uma pesquisa será a soma desses, ou seja

$$M_v = M_e - E_a - E_s$$

onde

M_v = Medida verdadeira
M_e = Medida estimada
E_a = Erro amostral
E_s = Erro não amostral ou sistemático

Os erros amostrais ocorrem em função do número de elementos da amostra e do processo de seleção desses elementos. Os erros não-amostrais, ou erro sistemático, são todos os erros cometidos durante o processo de pesquisa que não são decorrentes do tamanho ou do processo de seleção da amostra. Há referência de que, embora o erro sistemático refira-se aos erros ocorridos durante todo o processo da pesquisa, a fase de coleta de dados é uma das maiores fontes de erros não amostrais [14].

Para a realização de trabalho utilizando-se de enquetes, algumas precauções devem ser tomadas, com o intuito de melhor assegurar a acurácia e a validade dos resultados.

Uma das mais importantes tarefas no processo de conduzir uma pesquisa é: ter noção do quanto os seus dados são válidos e confiáveis. A idéia de validação se refere a quanto o processo de medição está isento de erros amostrais e não amostrais; já a confiabilidade se refere a quanto o processo está isento de erros amostrais [14]. Com relação à estimação da validade - um conceito maior e mais abrangente do que a confiabilidade - , seria ideal a comparação do resultado da medição realizada com o valor real da variável nas população [2]; porém, observamos que dada a dificuldade de conhecer o valor real da variável na população, tal procedimento é proibitivo. Assim, procura-se estimar a validade das medidas usando um ou mais dos diversos métodos de estimar validade [12].

Já o objetivo da fase de análise dos dados é obter significado nos dados coletados [14], podendo-se verificar os resultados obtidos. Assim, há uma fase em que se realiza o refinamento dos dados brutos coletados, transformando-os em dados que permitam a realização de análises e suas interpretações. Essa fase se chama processamento dos dados, e compreende as subfases de:

a - verificação - onde se impõe um padrão mínimo de qualidade dos dados brutos, minimizando a ambigüidade. A forma de proceder é a inspeção, e, quando necessário, a correção de cada questionário [14].

b - Codificação – Mattar [14] coloca a codificação como um procedimento técnico na qual os dados são categorizados. Mattar [14] ainda lembra os dois procedimentos com relação à codificação das questões abertas: o primeiro, partindo-se de um esquema de codificação definido *a priori*, e a aplicação desse esquema para a codificação; o segundo, espera-se terminar a pesquisa de campo, para então construir o esquema de codificação - a partir da verificação das respostas. O primeiro tem a desvantagem de exigir que o pesquisador tenha um profundo conhecimento do tema a ser pesquisado, além de ser desnecessário, caso o pesquisador tenha podido cobrir todas as possibilidades de questões abertas - quando então a questão poderia ter sido construída como uma questão fechada. O segundo, normalmente, exige uma recodificação dos dados.

c- Digitação – tendência atual, dadas as facilidades de utilizar-se tabulação eletrônica – normalmente, em planilhas eletrônicas ou programas de computador específicos, como o *Statistical Package for Social Sciences* (SPSS), ou o *Le Sphinx*.

d - Tabulação - após a digitação, passa-se a utilizar algumas das facilidades da tabulação eletrônica já integradas no programa (histogramas, filtros, verificação de valores estranhos, verificação de consistência, consideração ou não de casos atípicos etc.), para possibilitar uma análise mais eficiente.

Experimentação

Por definição, experimentação é um método de pesquisa onde se manipulam uma ou mais variáveis independentes e posteriormente, analisam-se as conseqüências dessa manipulação; normalmente, os sujeitos de pesquisa são designados aleatoriamente a grupos chamados “experimentais” , conforme a afirmação de diversos autores [10][18] [19] [13]. Assim, parte-se do princípio que há características que podem ser manipuladas, e outras que não podem. Para que se possa verificar a influência de uma dessas características no(s) sujeito(s) de pesquisa em questão, deve-se ter o máximo de controle sobre as características do sujeito, ou mesmo sobre as condições ambientais. Por essa razão, esse tipo de pesquisa geralmente é realizada em um ambiente ou local onde se tem grande capacidade de controle sobre as variáveis: o laboratório.

Nem sempre há a necessidade de haver experimentação sem que se realize uma designação aleatória dos indivíduos – embora realizar esse tipo de trabalho com essa forma aleatória fará com que o experimento torne-se mais forte do que algum outro sem essa designação.

Um interessante aspecto realizado freqüentemente em pesquisa experimental é haver alguma forma de interpretar a ausência da atuação da mudança no fenômeno a ser observado. Utiliza-se, para tanto, um grupo ou indivíduo de controle. Para esse grupo não é aplicado o mesmo tratamento que aos outros, e é realizada uma comparação entre todos os grupos experimentais e o de controle, para identificar as mudanças ocorridas pela aplicação das técnicas de mudança nos grupos experimentais. De modo geral, pode-se acreditar mais em resultados obtidos em pesquisas experimentais do que nos obtidos por outros métodos, dada a competência e satisfação dos padrões e critérios científicos [10].

A força da pesquisa experimental se encontra no alto controle da situação experimental. Contudo, ainda como pontos fortes dessa metodologia encontramos a possibilidade de manipulação das variáveis de forma isolada, ou em conjunto, e a possibilidade de replicação dos experimentos. Como aspectos fracos podemos citar a artificialidade do experimento, e a dificuldade de efetuar uma generalização.

Já com relação ao caráter da pesquisa experimental, podemos generalizá-lo como explicativo [19] ou explanatório, já que busca estabelecer relações causa-efeito, ou definir como ou porquê ocorrem determinados fenômenos, sob certas condições – os chamados fatores [25].

O fato de que o sujeito de pesquisa – no caso da Ciência da Computação – é mais informação do que energia ou matéria não deve fazer diferença na aplicação do tradicional método científico. Para entender a natureza do processo de informação, cientistas da Computação devem observar o fenômeno, formular explanações e teorias, e testá-las. Há ainda inúmeras teorias não testadas. Por exemplo, programação funcional, programação orientada a objetos, programação distribuída e métodos formais são todos utilizados no intuito de melhorar a produtividade do programador, e/ou a qualidade do programa. Embora utilizados há mais de 20 anos, esses métodos ainda não foram sistematicamente testados, não obstante haver uma série de importantes afirmações já proferidas sobre essas tecnologias [23].

Cientistas que usam a experimentação testam afirmações ou predições teóricas contra a realidade. Uma comunidade aceita gradativamente uma teoria, se todos os fatos conhecidos em seu domínio podem ser deduzidos da teoria, se os fatos resistiram a todos os numerosos testes experimentais, esse eles predisseram corretamente novos fenômenos [23].

Ademais, testa-se afirmações obtidas em estudos do tipo exploratório, as quais podem ou não serem corroboradas. Um exemplo conhecido de uso de experimentação na Ciência da Computação é o experimento de Knight e Levenson [11], na qual se analisa a probabilidade ocorrência de falhas, ao se utilizar a técnica de Tolerância a Falhas chamada de multiversões de programas. A teoria convencional afirma que a probabilidade de um programa multiversão falhar seria o resultado do produto das falhas das versões individuais. Contudo, John Knigth e Nancy Levenson observaram, através de experimentos, que a probabilidade era significantemente muito maior. Em essência, o experimento invalidou a afirmação da teoria convencional, afirmando que as falhas em um programa multiversões eram estatisticamente independentes. Outro exemplo de experimentação sobre estudos exploratórios são as redes neurais artificiais [23].

Para executar o processo de experimentação, há determinados requisitos a serem cumpridos. Caso se queira conduzir um experimento "puro", há três requisitos[19]: um dos requisitos é a manipulação intencional de uma ou mais variáveis independentes. Partindo desse requisito, pode-se notar a necessidade de identificar, dentro das características de nosso objeto de estudo, quais são aquelas que são dependentes, e quais não o são, dentro do processo existente na ocorrência de um fenômeno. Ademais, deve ser documentada a forma de efetuar a medição dessa característica. Um outro detalhe a ser definido em um trabalho desse tipo é a especificação das condições as quais estão sendo aplicadas à(s) variável(is) em estudo.

Um segundo requisito é a medição do efeito da variável independente sobre a dependente. Normalmente, busca-se experimentos anteriores para verificar como foram manipuladas essas variáveis. É imprescindível que se avalie essa manipulação antes da condução do experimento [19].

O último requisito é a forma de controlar a situação experimental. De modo geral, se há a garantia de controlar todas as variáveis – por exemplo, em um laboratório -, tem-se, então, o terceiro requisito para realizar um experimento puro. Para o caso de haver controle, a comunidade científica costuma representar da forma indicada na figura a seguir.

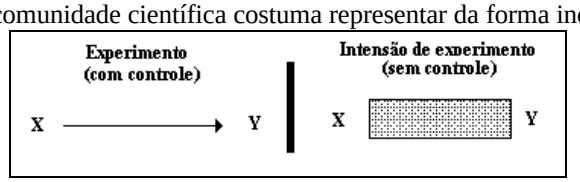

Figura 1. Indicação gráfica de controle de experimento.
Fonte: Adaptado de Sampieri et alli [19]

A figura acima apresenta a representação de controle – usando uma seta -, ou não – usando um retângulo preenchido. Outras ciências costumam definir construtos na forma de esquema, para explicar a metodologia aplicada no estudo. Assim, no caso desse tipo de experimento – chamado de "puro", ou tradicional -, o desenho do construto seria o seguinte:

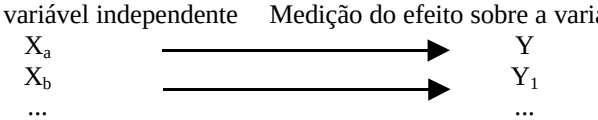

Figura 2. Desenho de experimento do tipo "puro".
Fonte: Traduzido de Sampieri et alli [19]

No caso acima, as letras “a”, “b”, etc. são indicativos de diferentes níveis de variação da variável independente. É importante indicar o sentido de crescimento do índice dos subscritos – por exemplo, “a” tem a máxima taxa de ocupação de memória, “n” tem a menor taxa de ocupação de memória.

Uma das mais importantes características para qualquer experimento é a repetibilidade [19][25]. Essa característica garante que os resultados podem ser verificados independentemente e, com isso, assegura confiança nos resultados – o que ajuda a eliminar erros, desvios, fraudes e mesmo trotes [23].

Com relação ao processo de manipular essas cariáveis, torna-se necessário – antes de realizar o experimento – definir como essas variáveis serão medidas. Será necessário avaliar o tempo de processamento? A taxa de ocupação de disco será maior? Qual o sistema operacional melhor indicado? Nem sempre se garante haver condições de controle elevado sobre as variáveis, ou sobre os procedimentos experimentais. Assim, podemos ter o tipo de experimentação chamado de quase-experimentos [19]. Algumas das formas de validação de experimentos são a utilização de vários grupos de comparação, buscar levantar equivalências entre os componentes dos grupos, e certificar-se de realizar a seleção de modo aleatório.

O principal elemento limitador da experimentação é o custo. De fato, há experimentos em que não há como minimizar o custo. Muitas vezes, no entanto, o custo pode ser elevado, mas dado o contexto, o custo pode vir ser a ser não proibitivamente caros – por exemplo, onde se envolvem vidas humanas. Experimentos com seres humanos envolvem desafios adicionais – grupos de controle, placebos, pré e pós-testes, balanceamento, estudos cegos e duplamente cegos e baterias de testes estatísticos.

Para o caso de não se poder efetuar a experimentação, pode-se realizar um estudo comparativo, ou uma simulação. O estudo comparativo – como o nome indica - busca efetuar comparações; o pesquisador deve, então, justificar o porquê escolheu determinadas variáveis para efetuar as comparações, assim como descrever as condições em que foram efetuadas as comparações. Estudos comparativos têm sido utilizados em diversas áreas, como reconhecimento de padrões, recuperação e armazenamento de informações, reuso de software, arquiteturas computacionais, avaliação de desempenho de componentes ou redes, análise numérica aplicada, algoritmos, robótica, compressão de dados, etc. A simulação é referida em capítulo posterior.

Outro fator limitador é ruído que pode ocorrer, em caso de não se considerar o elevado controle que o experimento deve ter, para sua correta condução. Talvez a maior prevenção que o cientista experimentador deve ter, seja a definição de variáveis e de suas formas de controle ANTES de realizar o experimento. Se o controle é impossível, os pesquisadores podem usar estudos de caso, estudos observacionais, ou outras técnicas investigativas.

Outros métodos utilizados em Computação

Diversas têm sido as variações e metodologias apresentadas em artigos de Ciências da Computação – de certa forma, bastante diversas daquelas tratadas até aqui. Zelkowitz e Wallace [25], em artigo publicado em 1998, propõem usar modelos experimentais para realizar uma atividade nem sempre presente nos artigos de eventos de Ciências da Computação: a validação. No mesmo artigo, fazem os autores algumas interessantes constatações, as quais tentar-se-á verificá-las no desenvolvimento do presente trabalho.

Embora voltado para a área de Engenharia de Software – e, de modo mais especial, para a experimentação -, há conceitos que podem vir a ser aplicados na metodologia de pesquisa em geral. Um deles é o tipo de método.

De acordo com Adrion [1], podemos ter quatro diferentes tipos de abordagens, a saber:

- Método científico – cientistas desenvolvem uma teoria para explicar um determinado fenômeno; eles propõem um(a) hipótese(s), e então testam variações da hipótese. Assim procedendo, eles coletam dados para verificar ou para refutar as afirmações da(s) hipótese(s).
- Método tecnológico – engenheiros desenvolvem e testam uma solução para uma hipótese. Baseado nos resultados dos testes, aquela solução é incrementada, até que não seja mais necessária nenhuma melhoria.
- Método empírico – um método estatístico é proposto como um meio para validar determinada hipótese. Diferente do método científico, pode não haver um modelo formal ou alguma teoria que descreva a hipótese. Os dados são coletados para verificar a hipótese.
- Método analítico – uma teoria formal é desenvolvida, e os resultados derivados daquela teoria podem ser comparados com observações empíricas.

Já Zelkowitz [25] busca categorizar modelos de validação de trabalhos – citando diversos métodos de pesquisa, tais como o estudo de caso, simulação e outros -, em três categorias: observacionais, históricos e controlados. Pode-se verificar que, normalmente, os métodos de pesquisa podem ser classificados baseados na possibilidade ou não de serem replicados, e da possibilidade de se exercer um maior controle ou não. Zelkowitz agrega, em se tratando de desenvolvimento de software, dois aspectos: a influência – ou impacto – que um projeto terá em um produto final, ou em um experimento; e propriedades temporais – considerando que a coleta de dados, para o trabalho de pesquisa ou de desenvolvimento de software poderá ser histórica ou atual.

Jenkins [9] classifica os métodos conforme estudos na área específica de Sistemas de Informação. Em seu estudo, apresenta 13 métodos; apresenta-se a seguir aqueles métodos com maior vínculo com a Ciência da Computação, a saber:

A - Modelagem ou demonstração matemática: dentre os métodos, certamente é o mais formal, buscando modelar o mundo real, e apresentar os resultados como resultados de equações matemáticas. Jenkins define como um “sistema

determinístico e fechado, na qual todas as variáveis – tanto as dependentes, quanto as independentes, são previamente conhecidas e consideradas no modelo". Observa-se, nesse método, que a intervenção não é possível.
B – Simulação experimental: método que usa um modelo fechado de simulação, para representar um segmento do mundo real. Aqui, os sujeitos humanos são expostos a esse modelo, e suas respostas são registradas. Jenkins coloca que "o pesquisador é quem determina o tempo e a natureza dos eventos experimentais.
C – Experimento de laboratório: é a experimentação tradicional, aqui destacando-se o uso de um ambiente de controle: o laboratório.
D – Simulação livre: método similar à simulação experimental, mas com a diferença de que o controle temporal e a natureza dos eventos não apenas são definidos pelo pesquisador, mas também pelo comportamento do objeto de pesquisa.
E – Experimento de campo: ao invés de ocorrer em um ambiente de total controle, o pesquisador usa o ambiente natural, "manipulando as variáveis independentes enquanto tenta controlar as mais importantes variáveis intervenientes, para então medir esses efeitos".
F – Experimento adaptativo: é o método de quase-experimento, que envolve medições antes e depois do experimento, além de necessitar de um grupo de controle, para efetuar as comparações que tornarão possíveis as medições.
G – Estudo de campo: similar ao experimento de campo, com a diferença que o pesquisador não manipula as variáveis independentes, mas unicamente as dependentes.
Os outros métodos apresentados por Jenkins são o estudo de caso – conforme já tratado em capítulo anterior -, e métodos com uma enorme ênfase na pesquisa social, tais como Análise de *feedback* de grupo, Pesquisa de opinião, Pesquisa ou observação participativa, Pesquisa em arquivos e Pesquisa filosófica.
Um método bastante recente e freqüentemente utilizado no Brasil é chamado de "dissertação-projeto" decorrente do avanço das áreas das Ciências Exatas e Tecnológicas, tendo sido colocada mais como uma metodologia de pesquisa tecnológica do que científica. Similar à intervenção, busca identificar um problema dentro de alguma área, caracterizar a área e desenvolver uma solução para o problema. Muitas vezes, esse problema é identificado apenas conceitualmente. A solução é, na maior parte dos casos, uma implementação de um programa de computador. Diferentemente da intervenção, ocorre que não se identifica, em momento algum, a influência do pesquisador-desenvolvedor no objeto a ser desenvolvido – mas unicamente com o problema de pesquisa a ser resolvido.

Considerações gerais sobre os métodos

A Ciência da Computação pode ainda ser considerada uma Ciência híbrida, uma vez que envolve aspectos técnicos – alguns com profunda fundamentação teórica, como Análise Computacional, Criptografia ou teoria da Computação -, educacionais, e organizacionais, entre outros.
Por esse ângulo, deve-se realizar cuidadosa análise ao procurar identificar o método empregado em um artigo ou outra publicação de caráter científico. Por exemplo, pode ser necessário identificar quais foram as condições nas quais aquelas informações foram proferidas; da mesma forma, se é possível efetuar repetições dos procedimentos ou não; o quão imparcial foi o pesquisador; qual é a experiência do pesquisador; como foram realizadas as análises; que hipóteses foram feitas; se é um primeiro estudo.

O estudo realizado

O método empregado foi a pesquisa survey, cujo instrumento foram os próprios anais de eventos. Para a definição dos temas componentes da sub-área escolhida – Segurança Computacional -, iniciamos um estudo bibliométrico sobre fontes bibliográficas altamente relevantes, identificando os temas através das palavras-chave encontradas nessas fontes.
Sobre anais de Congressos, Simpósios, *Workshops* e Encontros (principalmente, CRYPTO, EUROCRYPT, SAFECOMP, IFIP/SEC e ICICS) cujos temas tratavam sobre essas palavras-chave, iniciamos o processo de análise, identificando - no corpo dos artigos – características que identifiquem os métodos de pesquisa que foram utilizados para que se pudessem ser feitas as afirmações ou conclusões de pesquisa. Posteriormente, passou-se a montar os tópicos de interesse – as variáveis - em uma planilha eletrônica. Todos os eventos encontrados foram analisados e catalogados. Em algumas situações, no caso de ocorrência de propor algo, ou idéia inovadora, foi colocado como "Proposta de Conceito". As variáveis de interesse foram: dados identificatórios do artigo e do(s) autor(es), ano de publicação, evento, tema de pesquisa, tipo de método empregado, se esse método está explícito ou não, caráter da pesquisa, instituição e país de origem. Na prática, as principais variáveis compõe escalas do tipo nominal – o que limitou bastante o tipo de análises possíveis de serem realizadas. Ainda assim, foi possível, através de recursos de filtragem, realizar algumas análises.
Assim sendo, os métodos que se esperavam poderem aparecer são os indicados no quadro a seguir.

Método de pesquisa
Demonstração matemática
Dissertação-Projeto
Estudo comparativo
Estudo de campo
Estudo de caso
Experimentação
Histórica
Proposta de Conceito
Simulação
Survey

Quadro 1. Métodos de pesquisa esperados.
Fonte: Elaborado pelo autor.

Tabela 1. Percentuais encontrados.

Tipo de Método	% encontrado
Proposta de Conceito	42,33
Demonstração Matemática	42,22
Experimentação	5,04
Dissertação-Projeto	3,99
Estudo Comparativo	2,31
Estudo de Caso	1,05
Survey	0,52
Simulação	0,42
Não definido	0,31
Definição de Protocolo	0,31
Histórica	0,31
Bibliográfico	0,21

Fonte: Elaborado pelo autor

Com relação aos temas de pesquisa, os mesmos foram armazenados, resultando em 97 temas ou áreas de interesse, dentro da Criptografia – tendo muitos itens em comum com Segurança Computacional.

Resultados obtidos

Lembra-se que toda e qualquer informação afirmada no corpo do presente estudo refere-se ao universo definido – ou seja, métodos utilizados em artigos publicados em eventos internacionais da área de Segurança Computacional, com ênfase em Criptografia. Como em qualquer estudo realizado onde se utiliza a pesquisa *survey*, a generalização é possível, dentro de determinadas condições, e apenas para o universo considerado.

A seguir, apresentamos os resultados da pesquisa univariada sobre algumas das variáveis definidas anteriormente. As variáveis que eram de interesse para o presente trabalho eram o tipo de método empregado no artigo, o país onde a pesquisa foi realizada, a Instituição de pesquisa, o caráter da pesquisa, os temas mais pesquisados, e se o método de pesquisa era expresso de forma explícita, seja no *abstract*, seja no corpo do artigo.

Conforme se pode observar no gráfico da figura acima, os tipos preponderantes de metodologias que observamos são a Proposta de Conceito, seguido da Demonstração Matemática, da Experimentação e da Dissertação-Projeto. No caso dos dados acima analisados, o termo "indefinida" se refere ao fato de não apenas não se enquadrar em nenhuma das outras metodologias, mas também pelo fato de que esses artigos não traziam contribuição científica, mas meramente opiniões.

Com relação aos países de origem, observamos algumas especiais considerações. Como o intervalo temporal da coleta dos dados referia-se a artigos desde o início dos anos 80, ocorreram algumas alterações na Geografia política mundial. Assim, embora alguns artigos se referissem a países como "União Soviética", "Iugoslávia", "Checoslováquia", "Alemanha Ocidental", "Alemanha Oriental", optou-se por adaptar à realidade vigente na época do presente trabalho – primeiro semestre do ano 2000.

Sendo assim destaca-se, na tabela a seguir, os 15 países com maior produção na área de Pesquisa em Segurança Computacional, apresentados na tabela a seguir, por ordem decrescente.

Tabela 2 - Produção por país de origem

Países	% Ocorrências
Estados Unidos	24
Austrália	10
Reino Unido	9
Alemanha	9
França	8
Japão	7
Suíça	3
Canadá	3
Bélgica	3
Israel	3

Fonte: Elaborado pelo autor, com base na pesquisa realizada.

Observa-se que a produção desses 10 países, dentre os 31 observados, representa 79% do total. Embora não apresentado acima, houve 15 artigos onde não foi explicitada a origem, nem foi possível identificar por intermédio da origem de seus autores ou Instituições de Pesquisa – por se tratar de autor único, e/ou não citar a Instituição de Pesquisa.

Com relação ao caráter da pesquisa, constatou-se um predomínio nos caráteres Descritivo (55,15 %) e Exploratório (43,91%). Pequenas ocorrências observadas em artigos com caráter explanatório – causa/efeito, ou origem/conseqüência. Alguns artigos não puderam ter o seu caráter identificado. Já com relação à forma de apresentar o trabalho, observa-se que maior parte da forma de apresentar a metodologia é explicita (52,52%) – normalmente, no resumo. Para o restante, identificamos como realmente foram realizados procedimentos que permitiam chegar a conclusões pela leitura do corpo do artigo. Essas foram consideradas como a forma "implícita" de apresentar a metodologia.

A seguir, apresentam-se os resultados obtidos com relação à produtividade dos autores – cita-se os quinze autores de maior produtividade e instituições.

Tabela 3. Os quinze autores mais produtivos.

Autores mais produtivos	% ocorrências
MOTI YUNG	1,17
REIHANEH SAFAVI-NAINI	0,95
ED DAWSON	0,78
ROSS ANDERSON	0,73
BRUCE SCHNEIER	0,67
JACQUES STERN	0,67
JOVAN DJ. GOLIC	0,67
KOUICHI SAKURAI	0,61
TATSUAKI OKAMOTO	0,56
YULIANG ZHENG	0,56
COLIN BOYD	0,50
DAVID WAGNER	0,50
JOHN KELSEY	0,50
MIHIR BELLARE	0,50
YAIR FRANKEL	0,50

Fonte: Elaborado pelo autor, com base na pesquisa efetuada.

Tabela 4. As quinze instituições mais produtivas.

Instituições mais produtivas	% ocorrências
Queensland University of Technology	4,31
University of California	2,42
University of Wollongong	1,79
Katholieke Universiteit Leuven	1,68
MIT	1,47
University of London	1,37
Não declarada no do artigo	1,16
École Normale Superieure	1,16
Counterpane Systems	1,05
NTT Laboratories	1,05
AT&T LABS	0,95
University of Cambridge	0,95
Weizmann Institute	0,95
IBM T. J. Watson Research Center	0,84
Monash University	0,84

Fonte: Elaborado pelo autor, com base na pesquisa realizada.

A soma da produção desses autores é de quase 10 % da produção total. O número total de autores foi de 1792 pessoas. Um fato a ser destacado é a relação entre o número de pesquisadores e o número de artigos, o que resultou em uma razão de 1.88 pesquisadores por artigo. Em análise a ser apresentada posteriormente, observa-se que esse quadro está se modificando com o tempo – os artigos iniciais eram produções individuais, ao passo que atualmente nota-se a produção em equipes, mesmo encontrando-se em áreas geográficas distantes.

Uma das questões de interesse no trabalho era a identificação de quais são as instituições de pesquisa ou de ensino que têm apresentado maior produção científica. Essa informação pode denotar um interesse mais profundo – provavelmente, um grupo de pesquisa, caso mais de um autor tenha trabalhado esses temas, na mesma instituição -, ou mesmo a existência de um projeto de maior vulto do que publicações isoladas – identificado pela continuidade temporal da produção sobre o mesmo tema. A produção dessas instituições chega a aproximadamente 26 % da produção total pesquisada. Entre as cinco primeiras, duas são australianas, duas americanas e uma belga. Há autores que não citam informação alguma referente à instituição de origem – seja de ensino, seja de pesquisa ou mesmo comercial. Assim, foram agrupadas no item "Não declarada no corpo/resumo do artigo" – e corresponderam a um número total de ocorrências maior do que o número de ocorrências de algumas instituições de renome, tais como Weizmann Institute, Cambridge University, Siemens ou mesmo IBM. Uma informação de interesse é o fato de haver diversas instituições não acadêmicas, o que revela o interesse, a seriedade, e a preocupação comerciais do assunto de pesquisa – a Segurança Computacional (Criptografia).

O agrupamento de temas dos anais desses eventos totalizaram 97 temas, tendo ocorrido temas desde "agentes" até "vulnerabilidade", em ordem alfabética. Os temas dos primeiros trabalhos referiam-se, principalmente, a "esquemas de cifragem", tendo ocorrido outros mais recentemente – tais como "agentes", ou "dinheiro digital".

Esses quinze temas foram responsáveis por cerca de 48 % do total da produção, na pesquisa realizada. Observamos que há uma contribuição elevada de aspectos e temas relacionados à criptografia, por haver um maior número de eventos de onde foram colhidos os artigos - que consiste a maior parte da bibliografia encontrada em Anais de eventos na Biblioteca do Instituto de Informática da UFRGS.

A seguir, apresenta-se os temas de maior freqüência de aparecimento.

Tabela 5. Temas com maior ocorrência

Tema do artigo/pesquisa	% Ocorrências
CRIPTOANÁLISE	5,15
DINHEIRO DIGITAL	5,04
CIFRAS	4,73
ASSINATURAS DIGITAIS	4,31
CHAVE PÚBLICA	3,99
PROTOCOLOS	3,78
AUTENTICAÇÃO	3,26
ATAQUES/INVASÃO	2,84
CRIPTOGRAFIA	2,73
SEGURANÇA EM REDES	2,73
CONTROLE DE ACESSO	2,10
COMPARTILHAMENTO DE SEGREDO	2,00
COMÉRCIO ELETRÔNICO	1,89
COMPLEXIDADE	1,68
CURVAS ELÍPTICAS	1,58

Fonte: Elaborado pelo autor, com base na pesquisa realizada.

Limitações do estudo, considerações finais e perspectivas

A presente seção apresenta as dificuldades e considerações tomadas, algumas conclusões e sugestões para desenvolvimento de trabalhos futuros.

Dentre as dificuldades encontradas, podemos destacar problemas com a fonte – os anais dos eventos - e com os sujeitos de pesquisa. Ademais, a própria metodologia *survey* é orientada para enquetes e questionários, os quais têm – de modo geral -, diversos tipos de variáveis. As variáveis aqui definidas eram de escala nominal, o que limita o tipo de testes estatísticos possíveis de serem realizados. O processo de realizar a análise, utilizando-se da análise multivariada, ficou restrita ao uso e análise de tabelas de dados estratificados e/ou agrupados pelos mais diversos critérios.

Ademais, a bibliografia empregada – a qual continha o substrato para os sujeitos de pesquisa – foi muito concentrada em uma área específica: a Criptografia. Ademais, as edições que foram empregadas ficaram limitadas aos exemplares de 37 eventos ocorridos em diversos anos, em diversos locais. Houve uma concentração de exemplares dos últimos sete anos – havendo alguns exemplares com mais de 15 anos, nos quais se observa uma grande diferença nos enfoques nos quais se trabalhavam os conceitos de Criptografia, e mesmo na forma de escrita dos artigos

Uma das questões indiretas de interesse era o quão recentes eram os artigos utilizados em referências bibliográficas nos artigos que foram sujeitos de pesquisa. Essa questão poderia indicar, por exemplo, se um tema se baseia principalmente em artigos clássicos, ou se estaria aberto a novas idéias e temas. Para tanto, foi acrescentada uma coluna na planilha eletrônica, com o título ano da referência mais recente. Houve grande dificuldade em preencher essa coluna, uma vez que para referências de características semelhantes, os pesquisadores referenciavam de modo diferente. Isso ocorre provavelmente porque a forma de referenciar, na Ciência da Computação, é diferente de todas as outras Ciências. A mesma referência pode ser vista diferentemente em artigos diversos, conforme o exemplo abaixo:

> SAMPIERI, Roberto Hernández, COLLADO, Carlos Fernández, e LUCIO, Pilar Baptista. **Metodología de la Investigación**. México: McGraw-Hill, 1991. 514 p.
>
> ou
>
> [19] SAMPIERI, Roberto Hernández, COLLADO, Carlos Fernández, e LUCIO, Pilar Baptista. **Metodología de la Investigación**. México: McGraw-Hill, 1991. 514 p.

Essas formas diferentes, em um artigo com muitas referências, dificultam a identificação da referência mais recente. Destacamos que a primeira forma é a mais utilizadas em outras Ciências, sendo a segunda tradicionalmente utilizada – unicamente - em Ciência da Computação. Basicamente, pode-se identificar se um pesquisador é ou não da área pela forma de referenciar um artigo.

Uma observação interessante é que, em diversos artigos recentes, têm-se utilizado uma forma diferente de referência a múltiplos autores, conforme vê-se no exemplo abaixo:

> [SCL 91] SAMPIERI, Roberto Hernández, COLLADO, Carlos Fernández, e LUCIO, Pilar Baptista. **Metodología de la Investigación**. México: McGraw-Hill, 1991. 514 p.

Observa-se que o indicador inicial apresenta as iniciais de cada um dos três autores, ao invés das três letras iniciais do último sobrenome. A tendência observada em publicações pela IEEE é a seguinte:

[33] SAMPIERI, Roberto Hernández, COLLADO, Carlos Fernández, e LUCIO, Pilar Baptista. **Metodología de la Investigación**. México: McGraw-Hill, 1991. 514 p.

Sendo a trigésima terceira referência observada no artigo em questão.

Segurança computacional – em especial, a Criptografia - ainda apresenta dificuldades, graças à grande diversidade de problemas e de conhecimento. Por exemplo, há casos de temas de Criptografia em que se realiza pesquisa estritamente formal - como demonstrar matematicamente as afirmações de uma pesquisa envolvendo criptografia -, e também experimentos práticos, em ambiente real – como um protocolo criptográfico para garantir autenticidade de em remetente de uma conta bancária. Em ambos os casos, está sendo realizado um trabalho de pesquisa – porém de formas, conteúdos e objetivos diferentes.

Além disso, se considerarmos a área geral de Segurança Computacional, pode-se observar quebras de paradigmas realizadas com velocidade surpreendentemente dinâmica – podemos citar o caso de vírus de computador, os quais sempre se caracterizavam por limitar-se a apenas uma plataforma (ou PC, ou Mac) e, posteriormente, os vírus de macro vieram a mudar essa afirmação. No caso da Criptografia, têm-se observado grande ênfase para tópicos como curvas elípticas, criptografia quântica, criptografia biométrica e alguns protocolos criptográficos considerados esotéricos [20].

Outra dificuldade diz respeita aos conceitos empregados: há, mesmo na literatura especializada, alguma confusão no que tange a diversos termos. O mais crítico é usar a expressão "metodologia" – que seria o estudo dos métodos das Ciências -, com "método" – modo ou forma de proceder, para chegar a um fim.

Ainda há pouca definição em especificação de métodos de pesquisa na área estudada. Walter F. Tichy [22] realizou uma pesquisa *survey* sobre 400 artigos que afirmavam ter-se realizado experimentação; grande parte dos artigos (40 %) havia realizado "avaliação empírica", com nenhum suporte científico. Alguns artigos foram excluídos da amostra a ser analisada, em razão de se apoiarem apenas em demonstração matemática de teoremas - o que não pode ser provado por experimentos. Já Marvin Zelkowitz [25] realizou outra pesquisa por enquetes, envolvendo 600 artigos publicados – tendo sido apresentados como "foi aplicado experimentação" -, tendo observado que a maioria não validava experimentação específica da metodologia, e mesmo nenhuma; alguns autores utilizaram conceitos de validação de estudo de caso; e a terminologia de experimentação era, na maioria dos casos, descuidada. Destaca ainda que, aparentemente, o número de artigos sem validação parece aumentar.

Pouca surpresa ocorreu, ao agruparmos os temas de pesquisa por métodos, ou mesmo por países. No caso do agrupamento de temas por países, há a supremacia norte-americana em praticamente todos os temas de pesquisa. Já no caso de agrupamento de temas por métodos, os métodos mais formais foram observados em temas de pesquisa que realmente requeriam formalidade matemática; da mesma forma, métodos experimentais foram observados em temas recentes, os quais necessitam compor um corpo de conhecimento – ou seja, exatamente onde era necessário um estudo exploratório.

É relevante destacar o método de pesquisa, ao se escrever um artigo referente a um trabalho de pesquisa na área de Segurança Computacional? Aparentemente, sim. Qualquer produção onde se coloque de forma explícita o método utilizado permitirá que outros cientistas da Computação repliquem o estudo, ou que, pelo menos, tenham uma melhor condição de avaliar o processo de desenvolvimento e os resultados do estudo, bem como a sua qualidade.

Um aspecto que poderia vir a incrementar a qualidade do presente estudo seria o uso de outras formas de validação. Mesmo no caso da dissertação-projeto, metodologia que tem-se destacado – amplamente utilizada em países onde a Ciência da Computação é mais recente -, deve-se usar alguma forma de validação do trabalho efetuado, bem como explicitar todas as fases do estudo realizado. Por exemplo, Zelkowitz [25] propõe diversos métodos para efetuar validação – embora mais dirigido a experimentações -, entre os quais buscou classificar nas categorias a seguir:

- observacional - os métodos observacionais coletam dados relevantes, da mesma forma que desenvolvimento de um projeto – há pouco controle sobre o desenvolvimento de projetos inovadores;
- histórico – a coleta de dados ocorre sobre projetos já completados, tal como em um projeto baseado em engenharia reversa – o dado já existe, sendo necessário apenas a sua análise; e
- controlado – usa múltiplas instâncias de uma observação para validação estatística dos resultados – onde se pode empregar, por exemplo, uma simulação.

O presente estudo utilizou-se apenas da validação aparente – confiando à experiência de especialistas. Porém, dadas as escalas de nosso estudo, análises formais não seriam de possível realização, e não seria possível uso de técnicas de pré ou pós testagem, em função de não existir instrumento para avaliação – questionários -, o qual poderia conter viéses ou erros.

Certamente, para a realização de trabalhos futuros, deve-se aumentar a amostra trabalhada – ou seja, analisar um número maior de artigos, sendo a situação ideal aplicar a TODOS os eventos da área de interesse, de TODOS os anos em que esses eventos ocorreram. É interessante não apenas certificar-se de cobrir todos os eventos anuais, mas também de todas as áreas da segurança computacional. O fato de não ocorrer essa aleatoriedade implica a citação dessa fraqueza, em cada afirmação referente a conclusões no trabalho presente.

Um trabalho que certamente seria frutífero é a formalização do método chamado dissertação–projeto. Há a necessidade de formalizar formas de obtenção de dados, procedimentos de trabalho e, principalmente, validação do mesmo. Essa método, bastante usado no Brasil, tem aplicabilidade em toda a Ciência da Computação. Porém, o âmbito desse estudo transcende um trabalho individual vindo a constituir, certamente, um trabalho de doutorado.

Dados e análises completos do trabalho encontram-se disponíveis na Internet, na área de download da página em *www.sinpro-rs.org.br/vinicius.gadis.ribeiro* , arquivo "*TI1 - Métodos de Pesquisa empregados em Segurança Computacional - Criptografia*".

Bibliografia

[1] ADRION, W. R. Research Methodology in Software Engineering: Summary of the Dagstuhl Workshop on Future Directions in Software Engineering. **SIGSoft Software Eng. Notes**. New York, ACM Press: v. 18, n.1, p. 36-37, 1993.

[2] BABBIE, Earl. **Survey Research Methods**. 2. ed. Belmont: Wadsworth: 1990. 398 p.

[3] BENBASAT, I., MOORE, G. Development of Measures for Studying Emerging Tecnologies. In: HAWAII INTERNATIONAL CONFERENCE ON SYSTEMS SCIENCES (HICSS), Jan. 1992. **Proceedings...** Los Alamos: IEEE Society Press, p. 315-324.

[4] BRYMAN, Alan; BURGESS, Robert. **Analyzing Quantitative Data**. London: Routledge, 1995. 236 p.

[5] FINK, Arlene. **How to analyze survey data**. Thousand Oaks: Sage, 1995. v.8.104 p. *The Survey Kit*. il.

[6] FINK, Arlene. **The survey handbook**. Thousand Oaks: Sage, 1995. v.1.136 p. *The Survey Kit*. il.

[7] GIL, Antônio C. **Como elaborar projetos de pesquisa**. 3. ed. São Paulo: Atlas, 1991. 159 p. il.

[8] HOPPEN, Norberto; LAPOINTE, Liette; MOREAU, Eliane. **Um Guia para Avaliação de Artigos de Pesquisa em Sistemas de Informação**. Porto Alegre: PPGA-UFRGS, 1996. 18 p. Série Documentos para estudo.
Disponível na Internet.
http://www.cesup.ufrgs.br/PPGA/read/artigo/guia a.htm 08nov 96

[9] JENKINS, A. Milton. Research methodologies and MIS research. **Research methods in information systems**. Amsterdam: North-Holland, 1985. 320 p. Trabalho apresentado no IFIP WG8.2 Colloquium, 1984, Manchester. il.

[10] KERLINGER. Fred N. **Metodologia da Pesquisa em Ciências Sociais:** um Tratamento Conceitual. São Paulo: EPU, 1980. 386 p. il.

[11] KNIGHT, John C.; LEVENSON, Nancy G. An experimental Evaluation of the Assumption of Independence in Multiversion Programming. **IEEE Trans. Software Eng.**, New York, p. 96-109, Jan 1986.

[12] LITWIN, Mark S. **How to measure survey reliability and validity**. *The Survey Kit*, no.7. Thousand Oaks: Sage, 1995. 90 p.

[13] MARCONI, Marina de Andrade, e LAKATOS, Eva Maria. **Técnicas de Pesquisa**. 2. ed. São Paulo: Atlas, 1990. 234 p. il.

[14] MATTAR, N. **Pesquisa de *Marketing***. 3. ed. São Paulo: Atlas, 1996. v.2. 248 p. il.

[15] ORLIKOWSKI, W., BAROUDI, J. Studying Information Technology in Organizations: Research Approaches and Assumptions. **Information Systems Research**. New York, v.2, n.1, p. 1-28, Aug. 1991.

[16] PETER, P. J. Construct Validity: A Review of Basic Issues and Marketing Practices. **Journal of Marketing Research**. p. 6-17. May 1981.

[17] PINSONNEAULT, A.; KRAEMER, K. Survey Research in Management Information Systems: An Assessment. **Journal of Management Information Systems**. New York, v.10 n. 2, p. 75-106, Fall 1993.

[18] RIBEIRO, Vinicius G. **Um estudo sobre os métodos de pesquisa utilizados em Segurança Computacional**. Porto Alegre: PPGC da UFRGS, 2000. 70 p. TI 916.
Disponível na Internet em
<http://www.sinpro-rs.org.br/vinicius.gadis.ribeiro>

[19] SAMPIERI, Roberto Hernández; COLLADO, Carlos Fernández; LUCIO, Pilar Baptista. **Metodología de la Investigación**. México: McGraw-Hill, 1991. 514 p. il.

[20] SCHNEIER, Bruce. **Applied Cryptography**. New York: John Wiley & Sons, 1995. 624 p. il.

[21] STRAUB, D. Validating research instruments. **MIS Quarterly**. Minneapolis, v. 13, n.3, p. 147-169, Jun 1989.

[22] TICHY, Walter et al. Experimental Evaliation in Computer Science: A Quantitative Study. **J. Systems and Software**, New York, p. 1-18, Jan. 1995.

[23] TICHY, Walter F. Should computer scientists experiment more? **IEEE Computer**. New York, v. 15, n. 3,. p. 32-40, May 1998. il.

[24] YIN, Robert K. **Case study research:** design and methods.2. ed. London: Sage, 1994. 174 p. il.

[25] ZELKOWITZ, Marvin V.; WALLACE, Dolores R. Experimental models for validating technology. **IEEE Computer**. New York, v. 15, n. 3, p. 23-31, May 1998. il.