

CRIMES CIBERNÉTICOS

CPI

Pablo Ximenes

Information Security Research Team (INSERT) – Universidade Estadual do Ceará
Empresa de Tecnologia da Informação do Ceará (ETICE) – Governo do Ceará

Dados, Dados, Dados

- EM 1983, uma pesquisa sobre as finanças de consumidores nos EUA feita pelo Federal Reserve inflou erroneamente as estimativas de riqueza das famílias dos EUA em 1 trilhão e representou um incremento de 10% na estimativa.
- Isso foi resultado de uma única resposta incorreta de uma única pessoa!
- **Referência:** R. Avery, G. Elliehausen, and A. Kennickell. Measuring wealth with survey data: An evaluation of the 1983 survey of consumer finances. *Review of Income and Wealth*, 34(4):339–369, 1988.

Dados, Dados, Dados

- Em uma pesquisa de 2006 da FTC americana sobre Roubo de Identidade, duas respostas isoladamente representavam uma variação na estimativa da ordem de US\$37 bilhões.
- A adição dessas respostas no cômputo total multiplicaria as estimativas por três.
- **Referência:** Federal Trade Commission. Identity Theft Survey Report. 2007. www.ftc.gov/os/2007/11/SynovateFinalReportIDTheft2006.pdf.

Dados, Dados, Dados

- Em pesquisas sobre hábitos sexuais de homens e mulheres, os homens relatam um número de parceiras mulheres superior ao número de parceiros homens relatado pelas mulheres.
- Essa diferença varia de 3 a 9 vezes mais.
- Uma parcela ínfima dos homens exagera no número de parceiras (com exemplos que vão de 100 a 200). E esse “pequeno” exagero causa toda essa imprecisão absurda.
- **Referência:** M. Morris. Telling tails explain the discrepancy in sexual partner reports. Nature, 1993.

Problema 1: Concentração

- A forma mais comum de se estimar quantidades desconhecidas em uma população volumosa é através de pesquisas (entrevistas / survey).
- Quando o assunto são questões que se distribuem homogeneamente por toda a população a solução é simples: basta atingir um quantitativo que seja representativo da população.
- Exemplo: Todos tem a mesma chance de nascerem homens ou mulheres. Pontando essa característica é homogeneamente distribuída.

Problema 1: Concentração

- Quando se tratam de quantidades numéricas é comum haver concentração e desproporção.
- Por exemplo, se fizermos uma pesquisa da média salarial de Brasília e fizermos entrevistas apenas em residências no Plano Piloto, teremos um resultado que não representa a realidade.
- Onde entrevistar? Quantas pessoas de cada local? Qual o peso de cada resposta?
- Situações onde o que se investiga é distribuído de forma desigual exigem requintes técnicos muito mais complexos para se atingir um bom resultado.
- AMOSTRAS APARENTEMENTE REPRESENTATIVAS PODEM RESULTAR EM ESTIMATIVAS NÃO REPRESENTATIVAS. E resultam mesmo!

Problema 1: Concentração - Crimes Cibernéticos

- Em pesquisa do Gartner sobre Phishing, encontrou-se que a perda **média** por incidente é de US\$857, contudo a **mediana** da mesma pesquisa era de US\$200.
- Isso significa dizer que 59% das perdas se concentraram em 1% apenas dos entrevistados.
- Referência: Gartner. Phishing Survey. 2007.
<http://www.gartner.com/it/page.jsp?id=565125>.

Problema 1: Concentração - Crimes Cibernéticos

- A pesquisa do FTC relatada anteriormente, sofre do mesmo mal.
- Nela, a média foi de US\$1.876 e a mediana de US\$500, mesmo depois de excluídos os dois resultados "exagerados".
- Essa desproporção ocorreu também no levantamento de horas perdidas.

Problema 1: Concentração - Crimes Cibernéticos

- O Internet Crime Complaint Center (Integrado pelo próprio FBI, dentre outros), relatou em 2011 que a média para perdas monetárias de crimes cibernéticos era de US\$5.580 enquanto que a mediana era de US\$575.
- Isso é uma desproporção de 9.7 vezes mais.
- Isso significa que 1% da população concentra 78% de todo o dinheiro envolvido nas reclamações.

Problema 2: Respostas próprias sem verificação (self-reporting)

- Quantas calorias você consome? (Provavelmente mais do que você admite.)
- Quantos minutos de esteira você corre?
- Quantas mentiras você conta por dia?
- Etc.

Problema 2: Resposta próprias sem verificação (self-reporting)

- O que fazer quando a única fonte de informação pode ter conflito de interesses?
- Quem responde as pesquisas sobre Crimes Cibernéticos?
- Quem encomenda essas pesquisas?

Quanto custa o crime cibernético?

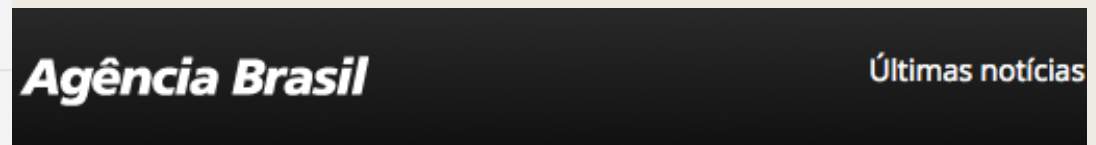
Symantec	McAfee
Era: US\$110 Bilhões	
US\$250 BILHÕES	US\$1 TRILHÃO

E no Brasil?



Economia

Pirataria e contrabando causam prejuízo de R\$ 30 bilhões para a indústria



mia

Contrabando dá prejuízo de R\$ 100 bilhões por ano ao Brasil

 Gosto < 10  Tweetar < 15  G+ < 0

URL: <http://agenciabrasil.ebc.com.br/econo>

E no Brasil?

- Alguns dados sobre a Pesquisa IPSOS/Oxford Economics de 2011 sobre a Pirataria em audiovisual no Brasil:

“Economic Consequences of Movie Piracy: Brazil” (Janeiro/2011)

- Pesquisa encomendada pelo Motion Picture Association of America (MPAA)

- As várias vezes que foi divulgada no Brasil, possuiu imprecisões:

- *Em artigo da International Intellectual Property Alliance (IIPA) menciona-se R\$7.3 bi em perdas econômicas*
- *Em entrevista do Sr. Ricardo Castanheira (diretor-geral da MPAA América Latina) ao Estadão, é citado R\$ 4 bilhões em perdas, com dados da mesma pesquisa.*

E no Brasil?

A metodologia comumente utilizada pelo IPSOS/Oxford nesses levantamentos:

- Entrevistas por Telefone, sem mecanismos para evitar a concentração
- Cálculo = (População Total do País) x (% de Pessoas que Declarou usar Conteúdo sem Pagar) x (Média de Filmes Pirateado por Pessoa)
- A metodologia abertamente reconhece que usa termos imprecisos e de significado mais abrangente que o escopo da pesquisa. Por exemplo, nunca usa a palavra “ilegal”, preferindo usar “sem autorização”.

Opinião da Academia acerca de Pesquisas sobre o Impacto Financeiro do Crime Cibernético

Our assessment of the quality of cyber-crime surveys is harsh: **they are so compromised and biased that no faith whatever can be placed in their findings.** We are not alone in this judgement. Most research teams who have looked at the survey data on cyber-crime have reached similarly negative conclusions.

“elas são tão mal-feitas e tendenciosas que não se pode colocar a menor fé em seus achados”

(Florencia, D. and Herley, C.; "Sex, lies and cyber-crime surveys", Workshop on Economics of Information Security, <http://research.microsoft.com/apps/pubs/default.aspx?id=149886>, June 2011.)

Por que se preocupar com os Dados?

- Eles pautam a nossa análise de Risco. Toda nova legislação traz ônus.
- Modelo de “Notice and Take Down” do DMCA.
 - *Coloca o Fardo nos Provedores, que por sua vez o transfere para usuários*
 - *Google recebeu mês passado 50.639.990 de pedidos Notice and Take Down*
 - *Pode ser usado para censura advinda de grupos econômicos*
 - *Onde termina o direito autoral?*
- **Criminalização de Ferramentas de Segurança**
 - *Pesquisa em segurança cibernética para no tempo. Exemplo: Falhas de Segurança*
 - *Não se pode punir o ferreiro por ter feito o martelo, por conta de um assassino que o usou. Exemplo: Livro sobre Vírus de Computador*

Apelo

- Nunca confiar cegamente em pesquisas.
- Exigir que a metodologia seja debatida e criticada. Seja posta à prova!
- Contar com as universidades brasileiras (temos menos conflitos de interesse!)
- Pesquisas que não consideram a desigualdade natural na distribuição em casos onde existe concentração ou que não possuem verificabilidade dos dados levantados não podem pautar decisões. Elas distorcem a realidade!

Obrigado!

- pablo@ximen.es
- <http://ximen.es/>
- <http://twitter.com/pabloximenes>