ZROZUMIENIE STUDIÓW I BADAŃ PRZESTĘPSTW KOMPUTEROWYCH

Ta część zawiera wskazówki dotyczące krytycznego czytania wyników badań dotyczących przestępczości komputerowej. Będzie również ostrzegać projektantów instrumentów badawczych, którym może brakować formalnego szkolenia w zakresie projektowania ankiet i analiz potrzebnych do profesjonalnego wsparcia w opracowywaniu kwestionariuszy i analizie wyników.

WATOŚĆ STATYSTYCZNEJ BAZY WIEDZY

Specjaliści od bezpieczeństwa są często pytani o przestępstwa komputerowe; na przykład klienci chcą wiedzieć, kto atakuje, które systemy, jak często, za pomocą jakich metod. Pytania te są postrzegane jako ważne, ponieważ dotyczą strategii zarządzania ryzykiem; teoretycznie, aby oszacować odpowiedni poziom inwestycji w bezpieczeństwo, pomocne byłoby rzetelne zrozumienie prawdopodobieństwa różnych poziomów szkód. Najlepiej byłoby ocenić poziom ryzyka organizacji, oceniając doświadczenia innych organizacji o podobnych cechach systemowych i biznesowych. Takie porównania byłyby przydatne w analizie konkurencji i postępowaniach sądowych dotyczących standardów należytej staranności i staranności w ochronie majątku przedsiębiorstwa.

OGRANICZNIE NASZEJ WIEDZY NA TEMAT PRZESTĘPCZOŚCI KOMPUTEROWEJ

Niestety w obecnym stanie bezpieczeństwa informacji nikt nie może udzielić wiarygodnych odpowiedzi na takie pytania. Istnieją dwie podstawowe trudności uniemożliwiające nam opracowanie dokładnych statystyk tego rodzaju. Trudności te znane są jako problemy z ustaleniem.

WYKRYWANIE

Pierwszy problem polega na tym, że nieznana liczba przestępstw wszelkiego rodzaju jest niewykryta. Na przykład, nawet poza polem przestępczości komputerowej, nie wiemy, ile oszustw finansowych jest popełnianych. Nie wiemy, ponieważ niektóre z nich nie są wykrywane. Skąd wiemy, że nie są wykrywane? Bo niektóre oszustwa są wykrywane na długo po ich wystąpieniu. Podobnie przestępstwa komputerowe mogą nie zostać wykryte przez ich ofiary, ale mogą zostać zgłoszone przez sprawców. W przełomowej serii testów w Departamencie Obrony (DoD) Agencja Obrony Informacji Systemowych stwierdziła, że bardzo niewiele z jej penetracji zaprojektowało wydaje się, że systemy niesklasyfikowane w DoD zostały wykryte przez menedżerów systemu. Badania te przeprowadzono w latach 1994–1996 i zaatakowały 38 000 systemów. Udało się około dwóch trzecich ataków; jednak wykryto tylko 4 procent tych ataków.1 Powszechnie uważa się, że w społeczności bezpieczeństwa informacji wykryto tylko jedną dziesiątą wszystkich przestępstw popełnionych przeciwko systemom komputerowym i korzystających z nich.

RAPORTOWANIE

Drugim problemem stwierdzenia jest to, że nawet w przypadku wykrycia ataków niewiele wydaje się być zgłaszane w sposób umożliwiający systematyczne gromadzenie danych. To powszechne przekonanie opiera się częściowo na niekwestionowanym doświadczeniu specjalistów ds. Bezpieczeństwa informacji, którzy przeprowadzili wywiady ze swoimi klientami; okazuje się, że ujawniło się tylko około 10 procent ataków na systemy komputerowe w takich wywiadach kiedykolwiek zgłaszano jakiekolwiek autorytety lub opinii publicznej. Wspomniane wcześniej badania DoD były zgodne z tym przekonaniem; z kilku wykrytych penetracji tylko ułamek 1% został zgłoszony odpowiednim władzom. Biorąc pod uwagę te problemy, statystyki dotyczące przestępczości komputerowej powinny być traktowane ze sceptycyzmem.

OGRANICZENIA DOTYCZĄCE STOSOWANIA STATYSTYK DOTYCZĄCYCH PRZESTĘPSTW KOMPUTEROWYCH

Uogólnienia w tej dziedzinie są trudne do uzasadnienia. Nawet gdybyśmy wiedzieli więcej o rodzajach przestępców i stosowanych przez nich metodach, nadal trudno byłoby uzyskać statystyki aktuarialne powszechne w dziedzinie ubezpieczeń. Na przykład ustanowienie jednolitych przepisów budowlanych w latach trzydziestych w Stanach Zjednoczonych doprowadziło do wzrostu ubezpieczenia od ognia jako rentownej firmy. Z oficjalnymi zapisami pożarów w budynki, które można opisać przy użyciu standardowej typologii, informacje statystyczne zaczęły zapewniać aktuarialną podstawę do wykorzystania prawdopodobieństwa pożarów i powiązanych kosztów do obliczenia rozsądnych stawek ubezpieczeniowych. W przeciwieństwie do tego, nawet gdybyśmy mieli dostęp do dokładnych raportów, trudno byłoby dokonać znaczących uogólnień dotyczących luk w zabezpieczeniach i częstości udanych ataków w dziedzinie technologii informacyjnej. Używamy oszałamiającej różnorodności i wersji procesorów, systemów operacyjnych, zapór ogniowych, szyfrowania, oprogramowanie aplikacyjne, metody i nośniki kopii zapasowych, kanały komunikacyjne, identyfikacja, uwierzytelnianie, autoryzacja, podział na partycje i operacje. W jaki sposób moglibyśmy uogólnić na podstawie danych dotyczących ryzyka w (powiedzmy) sieci opartej na komputerach typu mainframe z wieloma systemami wirtualnymi (MVS) w instalacji wojskowej na rodzaje zagrożeń, na jakie narażony jest intranet oparty na systemie UNIX w korporacji przemysłowej lub na Windows Serwer sieci Web oparty na nowej technologii (NT) w środowisku uniwersyteckim? Jest tak wiele różnice między systemami, że gdybyśmy mieli ustanowić wielowymiarową tabelę analityczną, w której każda zmienna byłaby osią, wiele komórek prawdopodobnie nie zawierałoby lub zawierało tylko kilka przykładów. Takie rzadkie macierze są bardzo trudne do wykorzystania w budowaniu modeli statystycznych do celów predykcyjnych.

PODSTAWOWA METODOLOGIA BADAŃ

To nie jest sekcja o badaniach nad naukami społecznymi. Jednak wiele dyskusji na temat przestępczości komputerowej wydaje się przyjmować opublikowane raporty jako ewangelię, nawet jeśli te badania mogą nie mieć żadnej ważności. W tej krótkiej części przyjrzymy się niektórym podstawom projektowania badań, aby czytelnicy mogli ocenić, ile wiary można włożyć w badania nad przestępczością komputerową wyniki.

NIEKTÓRE PODSTAWY PROJEKTOWANIA I ANALIZY STATYSTYCZNEJ

Sposób, w jaki naukowiec lub reporter reprezentuje dane, może stanowić ogromną różnicę w wrażeniach czytelników.

STATYSTYKI OPISOWE

Załóżmy, że trzy firmy zgłosiły te straty z powodu penetracji swoich systemów komputerowych: 1 mln USD, 2 mln USD i 6 mln USD. Możemy opisać te wyniki na wiele sposobów. Na przykład możemy po prostu wymienić nieprzetworzone dane; jednak takie listy mogą stać się niedopuszczalnie długie, ponieważ zwiększy się liczba raportów i trudno będzie zrozumieć dane pierwotne. Moglibyśmy zdefiniować klasy takie jak „2 miliony lub mniej” i „ponad 2 miliony” i policzyć, ile zdarzeń wystąpiło w każdej klasie:

Klasa Freq

≤ 2 mln USD 2

> 2 mln USD 1

Alternatywnie, możemy zdefiniować klasy o lepszej ziarnistości jako <1M $, ≥ 1M $, ale <2M $, i tak dalej; taka tabela może wyglądać tak:

Klasa Freq

<1 mln $

≥ 1 mln USD i <2 mln USD 1

≥ 2 mln USD i <3 mln USD 1

≥ 3 mln USD i <4 mln USD 0

≥ 4 mln USD i <5 mln USD 0

≥ 5 mln USD i <6 mln USD 0

≥ 6 USD 1

Zwróć uwagę, jak definicja klas wpływa na postrzeganie wyników: Pierwsza tabela sprawia wrażenie, że wyniki są skupione wokół 2 milionów dolarów i nie zawierają informacji o górnych lub dolnych granicach.

LOKALIZACJA

Jednym z najbardziej oczywistych sposobów opisywania danych jest stwierdzenie, gdzie znajdują się w określonym wymiarze. Tendencja centralna naszych trzech oryginalnych danych (1 mln USD, 2 mln USD i 6 mln USD) może być reprezentowana na różne sposoby; na przykład dwa popularne środki to:

* Średnia arytmetyczna lub średnia = $ (1 + 2 + 6) M / 3 = 3 mln USD

* Mediana (środek posortowanej listy strat) = 2 mln USD

Zauważ, że gdybyśmy próbowali obliczyć średnią i medianę z pierwszej tabeli (z jej przybliżonymi klasami), otrzymalibyśmy błędną wartość. Takie statystyki powinny być obliczane na podstawie oryginalnych danych, a nie tabel zbiorczych.

DYSPERSJA

Innym aspektem naszych danych, którego często potrzebujemy, jest rozproszenie - to znaczy zmienność. Najprostszą miarą dyspersji jest zakres: różnica między najmniejszą a największą wartością, którą znaleźliśmy; w naszym przykładzie moglibyśmy powiedzieć, że zakres wynosił od 1 miliona do 6 milionów dolarów lub 5 milionów dolarów. Czasami zakres jest wyrażany jako procent średniej; wtedy powiedzielibyśmy, że zakres wynosił 5⁄3 = 1,6. . . lub 167 procent. Wariancja (σ2) tych danych jest średnią odchyleń kwadratowych ze średniej arytmetycznej; wariancja trzech liczb wynosiłaby σ2 = (1−3) 2 + (2−3) 2 + (6−3) 2⁄3 = (4 + 1 + 9) ⁄ 3 ≈ 4,67.

Pierwiastek kwadratowy wariancji (σ) nazywany jest odchyleniem standardowym i jest często używany do opisu dyspersji. W naszym przykładzie σ = 67 4,67 ≈ 2,16. Dyspersja jest szczególnie ważna, gdy porównujemy szacunki dotyczące informacji z różnych grup. Im większa wariancja miary, tym trudniej jest utworzyć rzetelne uogólnienia dotyczące leżącego u podstaw zjawiska, jak opisano w następnej sekcji.

WNIOSKOWANIE : PRZYKŁADOWE STATYSTYKI A STATYSTYKI POPULACJI

Możemy dokładnie opisać wszelkie dane za pomocą statystyk opisowych; pytanie brzmi, co wtedy zrobimy z tymi środkami. Zazwyczaj spodziewamy się rozszerzyć wyniki w próbce lub podzbiorze populacji, aby uogólnić na temat populacji. Na przykład, możemy próbować oszacować straty związane z przestępczością komputerową w organizacjach komercyjnych mających biura w Stanach Zjednoczonych i zatrudniających ponad 30 000 pracowników. A może nasza próba reprezentowałaby organizacje komercyjne z biurami w Stanach Zjednoczonych i zatrudniające ponad 30 000 pracowników i których pracownicy ochrony sieci byli gotowi odpowiedzieć na kwestionariusz ankiety. W takich przypadkach staramy się wywnioskować cechy populacji z charakterystyki próbki. Statystycy twierdzą, że próbujemy oszacować statystyki parametryczne, korzystając ze statystyk przykładowych. Na przykład szacujemy wariancję parametryczną (populacyjną) (zwykle oznaczoną $\sigma 2$), mnożąc wariancję próbki przez $n / (n - 1)$. Zatem powiedzielibyśmy, że oszacowanie wariancji parametrycznej (s2) w naszej próbce wynosiłoby $s2 = 4,67 * 3/2 = 7$. Oszacowanie parametrycznego odchylenia standardowego (s) wynosiłoby $s = \sqrt{7} \approx 2,65$.

TESTOWANIE HIPOTEZ

Innym rodzajem wnioskowania, które próbujemy uzyskać z danych, jest testowanie hipotez. Załóżmy na przykład, że interesowało nas, czy istnieje związek między obecnością lub brakiem zapór ogniowych a występowaniem penetracji systemu. Wypracowalibyśmy hipotezę (hipoteza zerowa, czasami reprezentowana jako H_0), że nie było związku między dwiema niezależnymi zmiennymi, penetracją i zaporami ogniowymi, i przetestowali tę hipotezę, wykonując test niezależności tych zmienne. W naszym przykładzie prosty test niezależności chi-kwadrat dałby statystykę testową $\chi 2$ [1] = 2,636. Jeśli w populacji badanych systemów naprawdę nie było związku między penetracją a zaporami ogniowymi, wartość parametryczna tego statystyka wynosiłaby zero. W naszym wyimaginowanym przykładzie możemy wykazać, że tak duża wartość (lub większa) $\chi 2$ [1] wystąpiłaby tylko w 10,4% próbek pobranych z populacji, w której ściany ogniowe nie miały wpływu na penetrację. Innymi słowy, gdybyśmy wzięli wiele próbek z populacji, w której obecność zapór ogniowych nie była powiązana przy każdej zmianie szybkości penetracji, zobaczylibyśmy około 10,4 procent tych próbek, które generują statystyki $\chi 2$ [1] tak duże lub większe niż 2,636. Statystycy zgodzili się na pewne konwencje decydujące o tym, czy statystyka testowa odbiega od wartości oczekiwanej w hipotezie zerowej, aby uzasadnić wnioskowanie, że hipoteza zerowa jest błędna. Ogólnie opisujemy prawdopodobieństwo, że wartość null hipoteza jest prawdziwa - często pokazywana jako p (H0) - w ten sposób:

* Gdy $p (H_0) > 0,05$, mówimy, że wyniki nie są istotne statystycznie (często oznaczone symbolami ns);

* Gdy $0,05 \geq p (H_0) > 0,01$, wyniki są opisane jako istotne statystycznie (często oznaczone symbolem $*$);

* Gdy $0.01 \geq p (H_0) > 0.001$, wyniki są opisane jako wysoce istotne statystycznie (często oznaczone symbolami $**$);

* Gdy $p (H_0) \leq 0,001$, wyniki są opisane jako niezwykle istotne statystycznie (często oznaczone symbolami $***$).

LOSOWE PRÓBKOWANIE , ODCHYLENIE I ZMIESZANE ZMIENNE

Najważniejszym elementem pobierania próbek jest losowość. Mówimy, że próbka jest losowa lub losowa, gdy każdy członek populacji, którą studiujemy, ma takie samo prawdopodobieństwo bycia wybranym. Gdy populacja jest zdefiniowana w jeden sposób, ale próbka jest rysowana nieprzypadkowo, próbka jest opisana jako stronnicza. Na przykład, jeśli badana populacja ma być, powiedzmy, wszystkimi firmami na całym świecie zatrudniającymi ponad 30 000 pracowników pełnoetatowych, ale próbowaliśmy głównie z takich firm w Stanach Zjednoczonych, próbka byłaby

tendencyjna wobec firm amerykańskich i Podobnie, gdybyśmy mieli studiować bezpieczeństwo we wszystkich firmach w Stanach Zjednoczonych z ponad 30 000 pełnoetatowych pracowników, ale próbowaliśmy tylko od tych firm, które chciały odpowiedzieć na ankietę bezpieczeństwa, bylibyśmy zagrożeni o nieobiektywnej próbce. W tym ostatnim przykładzie, obejmującym studiowanie tylko tych, którzy odpowiadają na ankietę, mówimy, że jesteśmy potencjalnie zakłócającymi zmiennymi: przyglądamy się badaniom ludzi, którzy odpowiadają na ankiety i mają nadzieję, że są reprezentatywni dla większej populacji ludzi z wszystkich firm w pożądanej populacji. Ale co, jeśli ludzie, którzy chcą odpowiedzieć, to ci, którzy mają lepsze bezpieczeństwo, a ci, którzy nie reagują, mają straszne bezpieczeństwo? Wtedy odpowiedź na ankietę jest mylona z jakością bezpieczeństwa, a nasza stronnicza próba może nas łatwo wprowadzić w błąd, przeceniając poziom bezpieczeństwa w pożądanej populacji. Innym przykładem tego, jak zmienne mogą być mylone, są porównania wyników badań przeprowadzonych w różnych latach. Jeśli w tych dwóch latach nie przeprowadzi się dokładnie tych samych wywiadów, możemy wprowadzać w błąd indywidualne różnice w odpowiedziach ze zmianami w czasie; chyba że dokładnie te same firmy są reprezentowane, możemy mieć zakłócające różnice między firmami ze zmianami w czasie; jeśli wydarzenia zewnętrzne doprowadziły ludzi aby być mniej lub bardziej skłonnym odpowiedzieć na pytania zgodnie z prawdą, możemy być zakłopotani chęcią reagowania zmianami w czasie. Jeśli ankiety są przeprowadzane z różnymi pytaniami lub wykorzystywane przez różne grupy badawcze, możemy wprowadzać w błąd zmiany metodologii ze zmianami w czasie.

LIMITY ZAUFANIA

Ponieważ losowe próbki różnią się naturalnie wokół statystyk parametrycznych (populacji), nie jest zbyt pomocne podawanie punktowego oszacowania wartości parametrycznej. Na przykład, jeśli przeczytamy, że średnie szkody z przestępstw komputerowych w ankiecie wynoszą 180 000 USD na incydent, co to oznacza o populacji oznaczać? Aby wyrazić naszą ufność w statystyce próby, obliczamy prawdopodobieństwo, że będziemy mieć rację, jeśli podamy przedziałową ocenę wartości populacji. Na przykład możemy stwierdzić, że mielibyśmy 95-procentowe prawdopodobieństwo, że mielibyśmy rację twierdząc, że średnie obrażenia wyniosły od 160 000 do 200 000 USD. W innej próbie możemy być w stanie zawęzić te 95-procentowe limity ufności do 175 000 USD i 185 000 USD. Ogólnie rzecz biorąc, im większa wielkość próby, tym węższe granice ufności będą dla poszczególnych statystyk. Obliczanie limitów ufności dla statystyk zależy od pewnych niezbędnych założeń:

* Losowe pobieranie próbek

* Znany rozkład błędów (zwykle rozkład normalny - czasami nazywany rozkładem Gaussa)

* Równa wariancja przy wszystkich wartościach pomiarów

Jeśli którekolwiek z tych założeń jest błędne, obliczone granice ufności dla naszych szacunków będą błędne; to znaczy będą wprowadzać w błąd. Istnieją testy tych założeń, które analitycy powinni przeprowadzić przed raportowaniem wyników; jeśli dane nie są zgodne z rozkładami błędów normalnych, czasami można zastosować transformacje normalizujące. W szczególności procenty nie są zgodne z rozkładem normalnym.

TABELE AWARYJNE

Tabele awaryjne. Jednym z najczęstszych błędów w zgłaszaniu wyników badań jest dostarczenie tylko części historii. Na przykład można przeczytać stwierdzenia takie jak: „Ponad 70 procent systemów bez zapór ogniowych zostało przenikniętych w ubiegłym roku”. Takie stwierdzenie może być prawdziwe, ale nie można go poprawnie interpretować, co oznacza, że systemy z zaporami ogniowymi były z

konieczności mniej lub bardziej podatne na penetracja niż systemy bez zapór ogniowych. Oświadczenie jest niekompletne; aby to zrozumieć, potrzebujemy drugiej części domniemanej tabeli kontyngencji - odsetka systemów z zaporami ogniowymi, które zostały spenetrowane w zeszłym roku - przed dokonaniem jakichkolwiek stwierdzeń na temat związku między zaporami ogniowymi a penetracjami. W obu przypadkach ktoś mógłby powiedzieć, że „penetrowano 70 procent systemów bez zapór ogniowych", ale implikacje będą radykalnie różne w obu zestawach danych. Bez znajomości prawej kolumny oryginalne stwierdzenie byłoby bez znaczenia.

POWIĄZANIE A ZWIĄZWK PRZYCZYNOWY

Kontynuując nasz przykład ze wskaźnikami penetracji, kolejnym błędem, którego często nie uczą się osoby nieprzeszkolone podczas badania informacji statystycznych, jest błędne powiązanie związku przyczynowego. Wyobraź sobie, że badanie wykazało, że niższy odsetek systemów z gaśnicami został spenetrowany niż systemy bez gaśnice i że ta różnica była istotna statystycznie. Czy taki wynik musiałby oznaczać, że gaśnice spowodowały zmniejszenie penetracji? Wiemy, że o wiele bardziej rozsądne jest założenie, że gaśnice zostały zainstalowane w organizacjach, których polityka bezpieczeństwa i polityka bezpieczeństwa były bardziej rozwinięte niż w organizacjach, w których nie zainstalowano gaśnic. W tym wyimaginowanym przykładzie gaśnice mogą nie mieć żadnego skutku przyczynowego, niezależnie od odporności na penetrację. Wynik ten ilustruje wpływ zmiennych zakłócających: obecność gaśnicy ze stanem świadomości bezpieczeństwa i polityki.

GRUPY KONTROLNE

Na koniec, aby zakończyć nasz przykład penetracji, jednym ze sposobów odróżnienia związku i związku przyczynowego jest kontrola zmiennych. Na przykład można zmierzyć stan świadomości i polityki bezpieczeństwa, a także obecność lub brak gaśnic i dokonać porównań tylko wśród grup o tym samym poziomie świadomości i polityki. Istnieją również techniki statystyczne do matematycznego kontrolowania różnic w takich zmiennych niezależnych.

TESTU A PRIORI KONTRA A POSTERIORI

Amatorzy lub początkujący czasami zapominają o zasadzie losowego pobierania próbek, która leży u podstaw wszystkich wnioskowań statystycznych. Żaden z testów hipotez ani obliczeń limitu ufności nie działa, jeśli próbka nie jest losowa. Na przykład, jeśli ktoś wędruje przez supermarket i zauważa, że jabłka Granny Smith wydają się być większe niż jabłka Macintosha, wybranie próbki - nawet losowej próbki - jabłek, które konkretnie doprowadziły do hipotezy, nie pozwoli na wiarygodne obliczenia prawdopodobieństwa że zastosowanie ma taką samą średnią wagę. Problem polega na tym, że te konkretne jabłka w ogóle nie byłyby próbkowane, gdyby obserwator nie został zmuszony do sformułowania hipotezy. Więc nawet jeśli konkretne porównanie statystyczne daje przykładową statystykę, która wydaje się mieć prawdopodobieństwo, powiedzmy, 0,001, nie można wiedzieć, jak próbkowanie odbiegało od przypadkowości. Zastosowanie testów statystycznych do danych po zauważeniu interesującej wartości opisowej, porównania lub trendu jest znane jako testowanie a posteriori. Sformułowanie hipotezy, uzyskanie losowej próbki oraz obliczenie statystyk i prawdopodobieństw zgodnie z założeniami tych statystyk i prawdopodobieństw jest znane jako testowanie a priori. Dobrym przykładem niebezpieczeństw związanych z testowaniem a posteriori jest niefortunny zwyczaj przeszukiwania sekwencji wyników, takich jak długie ciągi zgadnięć zebrane w testach studenckich zdolności paranormalnych i obliczanie wartości statystycznych na starannie wybranych podzbiorach łańcuchów. Te testy a posteriori są następnie przedstawiane tak, jakby były a priori i powodują wielkie zamieszanie i argumenty, takie jak: „Spójrz, chociaż ogólna proporcja poprawnych domysłów wynosiła (powiedzmy) 50,003 procent w tym <<bardzo dużej liczbie>> domyśla się, że istniała seria <znacznie mniejszej liczby> przypuszczeń, które były poprawne <dowolna wartość

większa niż 50%> w danym czasie! Prawdopodobieństwo takiego wyniku jest <bardzo mała liczba>. Dowodzi to, że istniał rzeczywisty efekt <niezależnie od tego, jak wyglądało leczenie>. ”Niestety, długa seria liczb może wytworzyć dowolny pożądany ciąg nieprzypadkowo; istnieją nawet testy znane jako testy uruchomieniowe, które mogą pomóc badaczowi ocenić nielosowość takich zdarzeń. W praktyce statystyka ustanowiła konwencję ograniczania szkodliwych skutków testowania a posteriori: należy użyć poziomu prawdopodobieństwa 0,001 jako odpowiednika minimalnego prawdopodobieństwa hipotezy zerowej. Ten zwyczaj sprawia, że znacznie mniejsze jest prawdopodobieństwo, że porównanie a posteriori skłoni użytkownika do zaakceptowania tego, co w rzeczywistości jest przypadkową zmiennością, która przykuła czyjeś oko. Najlepszym rozwiązaniem błędu wynikającego z testów a posteriori jest użycie zupełnie nowej próbki do porównania. W przykładzie Apple można zapytać kierownika sklepu o nowe, niezauważone i losowo wybrane partie obu typów zgłoszeń. Statystyki porównawcze byłyby wówczas wiarygodne i można by oczekiwać, że będą podążać za rozkładem parametrycznym leżącym u podstaw obliczeń prawdopodobieństwa hipotezy zerowej. Populacje, z których te jabłka zostały wybrane, nadal musiałyby być dokładnie określone. Czy populacje byłyby jabłkami w tym konkretnym sklepie? Dla tego konkretnego łańcucha? W tym konkretnym regionie kraju lub świata?

METODY BADAWCZE MAJĄCE ZASTOSOWANIE DO PRZESTĘPCZOŚCI KOMPUTEROWEJ

WYWIADY

Wywiady z osobami mogą być pouczające. Ogólnie rzecz biorąc, wywiady zapewniają bogactwo danych, które są niedostępne w inny sposób. Na przykład można poznać szczegóły dotyczące przestępstw komputerowych lub motywacji i technik stosowanych przez przestępców komputerowych. Wywiady mogą być ustrukturyzowane (przy użyciu precyzyjnych list pytań) lub nieuporządkowane (umożliwiając ankieterowi udzielenie odpowiedzi na nowe informacje zadając dodatkowe pytania do woli). Ankieterzy mogą robić notatki lub nagrywać wywiady w celu późniejszej transkrypcji słowo w słowo. W niestrukturalnych ankieterach wykwalifikowani ankieterzy mogą sondować odpowiedzi w celu wyjaśnienia niuansów znaczenia, które mogą zostać utracone przy użyciu bardziej brutalnych technik, takich jak ankiety. Techniki, takie jak analiza tematyczna, mogą ujawniać wzorce odpowiedzi, które można następnie zbadać za pomocą eksploracyjnej analizy danych.2 Analiza tematyczna jest techniką organizowania informacji nie ilościowych bez narzucania wcześniej istniejących ram danych; eksploracyjna analiza danych wykorzystuje techniki statystyczne w celu zidentyfikowania możliwie interesujących relacji, które można przetestować z niezależnie pozyskanymi danymi. Takie techniki eksploracyjne mogą poprawnie obejmować testowanie a posteriori, jak opisano w sekcji wcześniejszej, ale wyniki są wykorzystywane do zaproponowania dalszych badań, które mogą wykorzystać testy a priori dla najlepszego wykorzystania zasobów.

GRUPY FOKUSOWE

Grupy fokusowe są jak wywiady grupowe. Ogólnie rzecz biorąc, facylitator używa listy z góry określonych pytań i zachęca uczestników do swobodnej reakcji i interakcji. Często postępowanie jest filmowane zza lustra w jedną stronę w celu późniejszej szczegółowej analizy. Taka analiza może obejmować komunikację niewerbalną, taką jak mimika i inna mowa ciała, ponieważ uczestnicy mówią lub słuchają innych mówiących o konkretnych tematach.

ANKIETY

Ankiety polegają na proszeniu ludzi o udzielenie odpowiedzi na ustaloną serię pytań z listami dopuszczalnych odpowiedzi. Mogą być przeprowadzane osobiście lub poprzez dystrybucję i pobieranie kwestionariuszy przez telefon, pocztę, faks i e-mail. Niektóre kwestionariusze zostały opublikowane w

sieci. Krytyczną kwestią przy rozważaniu wiarygodności ankiet jest stronniczość autoselekcji - oczywisty problem, który wynika z wyników ankiet tylko osób, które zgodziły się uczestniczyć. Przed oparciem krytycznych decyzji na danych z ankiety warto dowiedzieć się, jaki był odsetek odpowiedzi; chociaż nie ma absolutów, generalnie mamy tendencję do ufania wynikom ankiety bardziej, gdy odsetek odpowiedzi jest wysoki. Niestety, odpowiedź stawki za ankiety telefoniczne są często niższe niż 10 procent; odsetek odpowiedzi na ankiety pocztowe i e-mailowe może być mniejszy niż 1 procent. W takich okolicznościach bardzo trudno jest ustalić przypadek losowego pobierania próbek, a wszystkie wyniki z badań o niskim wskaźniku odpowiedzi powinny być postrzegane jako wskazujące na zakres problemów lub doświadczeń respondentów, a nie jako wskaźniki statystyk populacji. Jeśli chodzi o ankiety internetowe, istnieją dwa typy ze statystycznego punktu widzenia: te, które używają silnej identyfikacji i uwierzytelniania, oraz te, które nie. Te, które nie są podatne na oszustwa, takie jak wielokrotne głosowanie przez te same osoby. Te, które zapewniają indywidualne uniwersalne lokalizatory zasobów (URL), aby ograniczyć głosowanie do jednego na osobę, mają jednak takie same problemy z nastawieniem na samokontrolę, jak każde inne badanie.

WALIDACJA INSTRUMENTU

Wywiady i inne metodologie badań w dziedzinie nauk społecznych mogą cierpieć z powodu systematycznej tendencji respondentów do kształtowania odpowiedzi w celu zadowolenia ankietera lub wyrażania opinii, które mogą być bliższe normie w każdej grupie, którą sami widzą. Tak więc, jeśli dobrze wiadomo, że każda organizacja powinien mieć plan ciągłości działania, niektórzy respondenci mogą błędnie przedstawić stan swoich planów ciągłości działania, aby wyglądały lepiej niż są w rzeczywistości. Ponadto instrumenty ankietowe mogą zniekształcać odpowiedzi poprzez frazowanie pytań w sposób stronniczy; na przykład pytanie „Czy Twoja firma ma zakończoną działalność plan ciągłości? ”może mieć bardziej dokładny wskaźnik odpowiedzi niż pytanie„ Czy Twoja firma spełnia standardy branżowe w zakresie wypełnionego planu ciągłości działania? ”To drugie pytanie nie jest neutralne i prawdopodobnie zwiększy odsetek odpowiedzi„ tak ”. Sekwencja odpowiedzi może powodować reakcje stronnicze; narażenie na pierwsze możliwe odpowiedzi może przypadkowo ustalić punkt odniesienia dla respondenta. Na przykład pytanie o wielkość infekcji wirusowych może zapytać:

Czy w ciągu ostatnich 12 miesięcy Twoja organizacja doświadczyła całkowitych strat spowodowanych infekcjami wirusowymi

(a) 1 milion USD lub więcej;

(b) mniej niż 1 milion USD, ale większy lub równy 100 000 USD;

(c) mniej niż 100 000 USD;

(d) wcale?

Aby przetestować stronniczość, projektant może stworzyć wersje instrumentu, w których te same informacje są uzyskiwane przy użyciu przeciwnej sekwencji odpowiedzi:

Czy w ciągu ostatnich 12 miesięcy Twoja organizacja doświadczyła całkowitych strat spowodowanych infekcjami wirusowymi

(a) wcale;

(b) mniej niż 100 000 USD;

(c) mniej niż 1 milion USD, ale większy lub równy 100 000 USD;

(d) 1 mln USD lub więcej?

Kolejność pytań może wpływać na odpowiedzi; po udzieleniu konkretnej odpowiedzi na pytanie respondent będzie miał tendencję do udzielania odpowiedzi na kolejne pytania dotyczące tego samego tematu zgodnie z pierwszą odpowiedzią w serii. Aby przetestować tego rodzaju stronniczość, projektant może tworzyć wersje instrumentu z pytaniami w różnych sekwencjach. Inna technika walidacji instrumentu wprowadza pytania bez ważnych odpowiedzi lub bezsensownym żargonem, aby sprawdzić, czy respondenci krytycznie myślą o każdym pytaniu lub po prostu udzielają odpowiedzi, która pojawia się w ich głowach. Na przykład jeden może wprowadzić bezsensowne pytanie: „Czy Twoja firma wykorzystuje metody wykrywania zakłóceń kwantowych w stanie ustalonym do wykrywania włamań?" w kwestionariuszu dotyczącym bezpieczeństwa i unieważnia wyniki respondentów, którzy odpowiedzą twierdząco na to i inne diagnostyczne pytania. Wreszcie, niezależna weryfikacja odpowiedzi dostarcza mocnych dowodów na to, że respondenci odpowiadają zgodnie z prawdą. Takie natarczywe dochodzenia są jednak rzadkie.

METAANALIZA. Czasami warto ocenić hipotezę na podstawie kilku badań. Możemy łączyć prawdopodobieństwa P tej samej hipotezy zerowej z prób k z wykorzystaniem formuły

X2 = −2Σ ln P, gdzie X2 jest dystrybuowane jako χ2 [2k], jeśli hipoteza zerowa jest prawdziwa we wszystkich próbach. Załóżmy na przykład, że specjalista kryminalistyczny ocenia możliwość, że pliki dziennika zostały naruszone w określonym czasie.

* Prowadzi test rozkładu częstotliwości poszczególnych cyfr w danych dla podejrzanego okresu, aby ocenić prawdopodobieństwo, że rozkład jest zgodny z zerową hipotezą losowości. P dla testu dwustronnego chi-kwadrat wynosi 0,072ns.

* Patrzy na średnią liczbę I / O na sekundę w rekordach dla okresu podejrzanego i porównuje dane z tą samą statystyką w okresie kontrolnym przy użyciu ANOVA; P dla testu dwustronnego wynosi 0,096ns.

W tym przypadku X2 = −2 Σ ln P X2 = −2 ∗ (ln 0,072 + ln 0,096) = −2 ∗ (- 2,63109 -2,34341) = 9,948992 z 4 stopniami swobody. P dla hipotezy zerowej wynosi 0,0413 ∗. Innymi słowy, szanse obserwowania wyników obu testów przez przypadek same, jeśli podejrzane dane były zgodne z surowymi danymi z danych porównawczych, są istotne statystycznie. Istnieje powód, by odrzucić hipotezę zerową: ktoś może bardzo dobrze manipulować plikami dziennika dla podejrzanego okresu. Ta technika podlega ograniczeniom. Większość metaanaliz wymaga obliczenia prawdopodobieństwa hipotezy zerowej dla pojedynczego ogona w tym samym kierunku; nie ma sensu łączenie prawdopodobieństw ze sprzecznych hipotez za pomocą dwustronnego prawdopodobieństwa. Jednakże, jak w powyższym przykładzie, jeśli hipotezy zerowe są spójne, nawet dwustronne prawdopodobieństwa mogą być użytecznie połączone. Inny problem jest bardziej systemowy. Istnieją poważne powody do obaw, że badacze i wydawcy czasami tłumią wyniki eksperymentalne, które nie odpowiadają ich oczekiwaniom lub pragnieniom. Takie tłumienie powoduje, że opublikowane wyniki wydają się wspierać pożądany wynik. Identyfikacja takiego tłumienia jest trudna. Jednym z podejść jest zbadanie dystrybucji opublikowanych danych i poszukaj wskazówek wykluczenia danych. Na przykład, jeśli rozkład częstotliwości nieprzetworzonych danych w opublikowanym raporcie pokazuje nagłe zanikanie danych w jednym kierunku (np. Jeśli rozkład częstotliwości wygląda jak normalna krzywa, z wyjątkiem tego, że lewa strona nagle spada do zera w pewnym punkcie), to publikacja może używać obciętego zestawu danych. Metaanaliza oparta na danych o wątpliwej trafności sama w sobie będzie wątpliwa. Syf na wejściu, syf na wyjściu.

PODSUMOWANIE

Podsumowując, wszystkie badania dotyczące przestępczości komputerowej powinny być dokładnie zbadane, zanim będziemy polegać na ich wynikach. Kilka podstawowych pytań na temat takich badań:

* Jaka jest populacja, z której próbujemy?

* Biorąc pod uwagę stronniczość autoselekcji, reprezentatywni dla szerszej populacji są respondenci, którzy zgodzili się wziąć udział w badaniu lub ankiecie?

* Jak duża jest próbka?

* Czy autorzy testują założenia losowości, normalności i równości wariancji przed zgłoszeniem miar statystycznych?

* Jakie są przedziały ufności dla zgłaszanych statystyk?

* Czy porównania zakłócają zmienne?

* Czy korelacje są błędnie interpretowane jako relacje przyczynowe?

* Czy instrumenty testowe zostały zatwierdzone?