Zeszyty Naukowe
Wyższej Szkoły Bankowej we Wrocławiu
Nr 20/2011

Agata Strzelczyk

Wyższa Szkoła Bankowa we Wrocławiu

Sposoby pomiaru ilości i wartości informacji

Streszczenie. Artykuł podejmuje tematykę pomiaru informacji. W pracy przedstawiono pojęcie informacji i omówiono sposoby pomiaru zarówno ilości, jak i wartości informacji.

Słowa kluczowe: informacja, ilość informacji, wartość informacji, miary wartości informacji, miary ilości informacji

Wstęp

Z pojęciem informacji można spotkać się na każdym kroku życia codziennego, informacja bowiem ma w dzisiejszym świecie niezwykłe znaczenie zarówno w życiu gospodarczym, społecznym, politycznym, jak i naukowym. Ludzie dążą do zdobywania jak najświeższych, najbardziej aktualnych, adekwatnych i dokładnych informacji. Gotowi są w tym celu na wiele poświęceń, często z narażeniem zdrowia czy życia włącznie. Jak zatem ocenić zdobytą informację? Jej ilość i wartość. Kiedy informacja *A* jest „lepsza” od informacji *B*? Celem tego artykułu jest przedstawienie koncepcji próbujących odpowiedzieć na te pytania.

Na wstępie jednak należy podkreślić, że pojęcie informacji nie zostało jednoznacznie zdefiniowane. Jest to termin interdyscyplinarny, stąd wiele definicji powstałych na użytek rożnych dziedzin nauki. Według *Nowej encyklopedii PWN* informacja to „w języku potocznym konstatacja stanu rzeczy, wiadomość; w dyscyplinach naukowych miara niepewności zajścia pewnego zdarzenia (otrzymania określonego wyniku pomiaru, wyemitowania określonej wiadomości przez źródło) spośród skończonego zbioru zdarzeń możliwych", ale także „obiekt abstrakcyjny, który w postaci zakodowanej (dane) może być przechowywany na nośniku danych, przesyłany (np. głosem, falą elektromagnetyczną, prądem elektrycznym), przetwarzany (w trakcie wykonywania algorytmu) i użyty do sterowania (np. komputerem steruje program będący zakodowaną informacją)"[1]. Według Głuszkowa informacja „są to wszelkie wiadomości o procesach i stanach dowolnej natury, które mogą być odbierane przez organy zmysłowe człowieka lub przez przyrodę"[2]. Niektórzy badacze, jak na przykład Wiener, definiowali informację poprzez negację, czyli stwierdzenie, czym informacja nie jest. Oto cytat: „Mechaniczny mózg nie wydziela myśli, jak wątroba wydziela żółć, zdaniem dawniejszych materialistów, ani też nie wydaje jej w postaci energii, jak robi to mięsień w swoim działaniu. Informacja jest informacją, a nie materią czy energią"[3].

Już powyższe przykłady pokazują, jak wiele trudności wynika z próby zdefiniowania informacji. Równie dużo problemów sprawia pomiar jej ilości czy też wartości.

1. Ilość informacji

Ilościowa teoria informacji, której początki sięgają pierwszej połowy XX w. podjęła próbę pomiaru ilości informacji. Prekursorzy tej teorii unikali jednak definiowania pojęcia informacji, poprzestając tylko na jej ilości. Pierwsza miara ilości informacji zaproponowana została przez R.V.L. Hartleya w 1928 r. Według tej koncepcji ilość informacji zawartej w zdarzeniu polegającym na rozpatrywaniu jednego z n obiektów należących do n-elementowego zbioru to:

$$I = \log n, \qquad (1)$$

gdzie I – ilość informacji.

[1] *Nowa encyklopedia PWN*, t.3, Wydawnictwo Naukowe PWN, Warszawa 1995, s. 53-54.

[2] A. Baborski, M. Duda, S. Forlicz, *Elementy cybernetyki ekonomicznej*, PWE, Warszawa 1977, s. 153.

[3] N. Wiener, *Cybernetics or Control and Communication in the Animal and the Machine*, New York, Wiley 1948, za: M. Mazur, *Jakościowa teoria informacji*, Wydawnictwa Naukowo-Techniczne, Warszawa, s. 19.

Zgodnie ze stwierdzeniem, że „w pojęciu informacji istotne jest nie samo zaistniałe zjawisko, lecz jego stosunek do zbioru zdarzeń, które mogły były zaistnieć”[4], koncepcja Hartleya została dopracowana i rozwinięta przez C.E. Shannona w 1948 r. Według niego, mierząc ilość informacji otrzymanej przy zajściu zdarzenia x_i niezbędne jest uwzględnienie także prawdopodobieństwa p_i zajścia tegoż zdarzenia, co przedstawić można za pomocą wzoru:

$$I(x_i) = -\log_m p_i, \tag{2}$$

gdzie:

$I(x_i)$ – ilość informacji otrzymanej przy zajściu zdarzenia x_i,
p_i – prawdopodobieństwo zajścia zdarzenia x_i,
m – podstawa logarytmu.

W zależności od przyjętej podstawy logarytmu we wzorze (2) miarą ilości informacji może być:

– **nit** lub **nat** (*natural units*), dla $m = e$,
– **dit** lub **hartley** (*decimal digits*), dla $m = 10$,
– **bit** (*binary digits*), dla $m = 2$,
– inna jednostka powstała na skutek przyjęcia jeszcze innej podstawy logarytmu.

Pomimo tej różnorodności najczęściej spotykaną jednostką miary ilości informacji jest bit. Komunikat dostarcza 1 bitu informacji wówczas, gdy nastąpiło jedno z dwóch, tak samo prawdopodobnych zdarzeń. A oto kilka przykładów:

– Rzucamy sześcienną kostką do gry. Niech zdarzenie A oznacza wypadnięcie parzystej liczby oczek, zdarzenie B – wypadnięcie nieparzystej liczby oczek, zdarzenie C – wypadnięcie liczby oczek podzielnej przez 3, a zdarzenie D – liczby oczek równej 1, 2, 3, 4, 5 lub 6. Jakich ilości informacji dostarcza komunikat o zajściu każdego ze zdarzeń?

  Ponieważ $p_A = p_B = 0{,}5$, więc $I(A) = I(B) = -\log_2\left(\frac{1}{2}\right) = 1\,\text{bit}$.

  Ponieważ $p_C = \frac{1}{3}$, więc $I(C) = -\log_2\left(\frac{1}{3}\right) \approx 1{,}58\,\text{bita}$.

  Ponieważ $p_D = 1$, więc $I(D) = -\log_2(1) = 0\,\text{bitów}$.

– Losujemy jedną kartę spośród talii 52 kart do gry. Niech zdarzenie A oznacza wylosowanie karty koloru czarnego, zdarzenie B wylosowanie karty koloru trefl, a zdarzenie C wylosowanie asa. Jakich ilości informacji dostarcza komunikat o zajściu każdego ze zdarzeń?

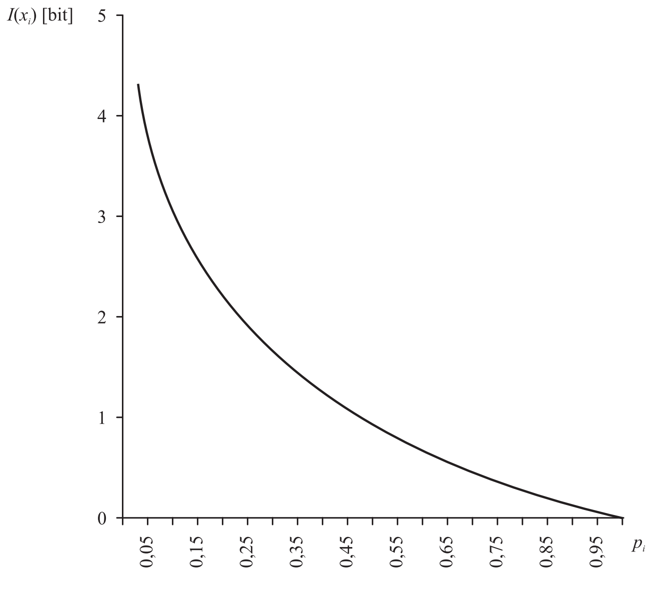
  Ponieważ $p_A = 0{,}5$, więc $I(A) = -\log_2\left(\frac{1}{2}\right) = 1\,\text{bit}$.

[4] W.M. Głuszkow za: J.L. Kulikowski, *Informacja i świat, w którym żyjemy*, Wiedza Powszechna, Warszawa 1978, s. 43.

Ponieważ $p_B = 0{,}25$, więc $I(B) = -\log_2\left(\frac{1}{4}\right) = 2$ bity.

Ponieważ $p_C = \frac{1}{13}$, więc $I(C) = -\log_2\left(\frac{1}{13}\right) \approx 3{,}7$ bita.

Powyższe przykłady pokazują, że jeśli $p_A < p_B$, to $I(A) > I(B)$. Oznacza to, że im mniejsze jest prawdopodobieństwo zajścia zdarzenia x_i, tym większa jest ilość informacji, którą niesie komunikat o zajściu tegoż zdarzenia. W skrajnej sytuacji – gdy prawdopodobieństwo zajścia zdarzenia x_i jest równe 1 – ilość informacji wynosi 0 (rys. 1). Fakt ten jest zgodny nie tylko ze wzorem (2), ale także z intuicją każdego człowieka. Im bardziej prawdopodobne jest zajście danego zdarzenia, tym mniejszą ilość informacji niesie. Dodatkowo komunikat, który dostarcza wiadomość znaną już wcześniej (prawdopodobieństwo równe 1), nie dostarcza żadnej informacji.



Rys. 1. Zależność ilości informacji otrzymanej przy zajściu zdarzenia x_i od prawdopodobieństwa zajścia zdarzenia x_i

Ź r ó d ł o: opracowanie własne.

Jak już wspomniano powyżej, w zależności od przyjętej podstawy logarytmu we wzorze (2) miary ilości informacji mogą być różne. Należy jednak pamiętać, że wykorzystując własność logarytmów:

$$\log_a x = \frac{1}{\log_b a} \cdot \log_b x \text{, gdzie } a, b > 1, \tag{3}$$

można dokonywać zamiany jednostek miar ilości informacji. Korzystając z (3) otrzymano:

$$\log_{10} x = \frac{1}{\log_2 10} \cdot \log_2 x = 0{,}301 \cdot \log_2 x,$$

$$\log_e x = \frac{1}{\log_2 e} \cdot \log_2 x = 0{,}693 \cdot \log_2 x.$$

Zatem:

1 dit = 0,301 bita, a 1 nit = 0,693 bita.

W tym miejscu pojawia się zatem pytanie: dlaczego najpopularniejszą miarą ilości informacji jest bit? Dlaczego najczęściej stosowany jest system binarny? W odpowiedzi przytoczony zostanie krótki przykład obrazujący, że system binarny jest najbardziej efektywnym systemem służącym do pomiaru informacji.

Założono, że marynarz ma za zadanie zasygnalizować liczby od 0 do 127, wykorzystując w tym celu flagi. Na początku decyduje, że każda liczba reprezentowana będzie za pomocą jednej flagi. Aby wykonać zadanie, potrzebuje zatem 128 flag. Rezygnuje z pomysłu i wybiera system dziesiętny. Wówczas niezbędnych jest 21 flag – 10 dla jednostek, 10 dla dziesiątek i 1 dla setki. Na koniec decyduje posłużyć się systemem binarnym. W tym przypadku potrzebuje tylko 14 flag. Siedem z nich to „0", a drugie 7 to „1". 14 flag zamiast pierwotnych 128 – czysta oszczędność[5].

Należy pamiętać, że na podstawie wzorów (1) i (2) określić można ilość informacji zawartą w komunikacie dotyczącym zdarzenia tylko dla zmiennych losowych o rozkładzie dyskretnym. W celu pomiaru ilości informacji wykorzystać można **entropię**. „Entropia informacyjna jest miarą nieokreśloności zdarzeń przy określonym stanie niewiedzy o nich"[6]. Ilość informacji obliczona za pomocą entropii, jako miary niepewności, to różnica pomiędzy entropią przed i po otrzymaniu informacji. Jeśli możliwe jest tylko jedno zdarzenie z prawdopodobieństwem 1, to wówczas entropia $H = 0$.

Dla zmiennej losowej o rozkładzie dyskretnym określona jest wzorem:

$$H = \sum_{i=1}^{n} p_i I(x_i) = -\sum_{i=1}^{n} p_i \log_2 p_i, \tag{4}$$

gdzie:

p_i – prawdopodobieństwo zajścia zdarzenia x_i,

[5] H.Ch. von Baeyer, *Information. The new language of science*, Harvard University Press, Cambridge, Massachusetts 2004, s. 31.

[6] S. Forlicz, *Informacja w biznesie*, PWE, Warszawa 2008, s. 22.

x_i – i-te zdarzenie traktowane jako i-ta realizacja zmiennej losowej X,
n – liczba możliwych stanów zmiennej losowej X.

W przypadku zmiennej losowej ciągłej określona jest za pomocą następującego wzoru:

$$H = -\int_{-\infty}^{\infty} f(x)\log_2 f(x)dx, \tag{5}$$

gdzie $f(x)$ – funkcja gęstości zmiennej losowej X.

Kolejną koncepcją wartą przedstawienia jest miara ilości informacji zaproponowana przez A. Kołmogorowa[7] w 1969 r. opisana wzorem:

$$I(a) = \min d(a), \tag{6}$$

gdzie:

$I(a)$ – ilość informacji,
a – algorytm należący do pewnej rodziny algorytmów A rozwiązujących dany rodzaj zadań,
$d(a)$ – miara długości algorytmu a (liczba poleceń, rozkazów, instrukcji w algorytmie), która w najprostszym przypadku równa będzie liczbie operacji wchodzących w skład algorytmu a.

Miara zaproponowana przez Kołmogorowa, w przeciwieństwie do omawianych wyżej, nie uwzględnia prawdopodobieństwa wystąpienia danego zdarzenia. Niestety nie uwzględnia także wewnętrznej struktury algorytmów. Wszystkie operacje będące częścią algorytmów uznaje za równe, bez względu na stopień ich złożoności.

Inną, niezwiązaną z prawdopodobieństwem zajścia zdarzeń, koncepcję pomiaru ilości informacji zaproponował w 1984 r. J. Hintikka. Przyjął, że:

$$I = L(i), \tag{7}$$

gdzie:

I – ilość informacji
$L(i)$ – ilość pracy wykonanej przez system na podstawie otrzymanej informacji.

Za podstawową jednostkę ilości informacji uznał taką ilość informacji, która potrzebna jest do wykonania jednego dżula pracy. Jest to sposób na zmierzenie na przykład ilości informacji, która zawarta jest w zadaniu przekazanym robotnikowi, polegającym na przeniesieniu określonego przedmiotu w określone miejsce[8].

---

[7] B. Stefanowicz, *Informacja*, Wydawnictwo SGH, Warszawa 2004, s. 85.
[8] Tamże, s. 86.

Sposobem pomiaru ilości informacji, która łączy w sobie koncepcje zarówno Hartleya, Shannona, jak i Kołmogorowa jest miara ilości informacji zaproponowana przez B. Stefanowicza[9], opisana wzorem:

$$\mu(K) = [\mu(O)^2 + \mu(A)^2 + \mu(a)^2]^{1/2}, \tag{8}$$

gdzie:

$\mu(K)$ – ilość informacji dostarczonej przez komunikat K,

$\mu(O)$ – ilość informacji, jaka wynika z uwzględnienia w komunikacie K pewnego konkretnego obiektu O; może być liczona za pomocą miary Hartleya, gdzie n będzie równe liczbie obiektów należących do zbioru, z którego pochodzi O,

$\mu(A)$ – ilość informacji, która wynika z uwzględnienia w komunikacie K atrybutu A; może być mierzona za pomocą miary Hartleya, gdzie n to liczba wszystkich atrybutów należących do zbioru zawierającego atrybut A,

$\mu(a)$ – ilość informacji zależna od wartości przyjmowanej przez atrybut A; w zależności od jej rodzaju, wyliczenia miary $\mu(a)$ można dokonać na różne sposoby:

- jeśli atrybut A jest opisem stanu obiektu O, a pojawienie się wartości a w komunikacie K wiąże się z odpowiednim prawdopodobieństwem, to ilość $\mu(a)$ wyliczyć można za pomocą miary Shannona (2),
- jeśli atrybut A oznacza strukturę obiektu O, a zaś jest opisem konkretnej struktury, to $\mu(a)$ wyliczyć można za pomocą miary Hartleya (1),
- jeśli A oznacza procedurę, a zaś jest pewną konkretną procedurą, to ilość $\mu(a)$ wyliczyć można za pomocą miary Kołmogorowa (6).

Uwaga: Aby zapewnić wewnętrzną spójność ilości informacji, podstawy logarytmów przy obliczaniu miar $\mu(0)$, $\mu(A)$ i $\mu(a)$ powinny być takie same. Pozwoli to na obliczenie wartości z uwzględnieniem odpowiedniej miary: bitu, nitu, ditu lub innej.

2. Wartość informacji

Jak już pokazano, wiele trudności sprawia znalezienie jednego uniwersalnego sposobu pomiaru ilości informacji. Niestety na podobne problemy napotkać można przy pomiarze wartości informacji. Wynika to z faktu, że wartość informacji rozumiana jest często w różny, zupełnie odmienny sposób.

---

[9] Tamże, s. 86-87.

Mierząc wartość informacji *ex ante* istnieje konieczność oszacowania wartości oczekiwanej korzyści, powstałych na skutek zdobycia danej informacji (co jest na ogół bardzo trudne).

W przypadku pomiaru wartości informacji *ex post* sposobów pomiaru może być kilka, w zależności od potrzeb czy też subiektywnych preferencji. Poniżej zostanie przedstawionych kilka z nich.

Miary ilości informacji. Z uwagi na fakt, że wartość informacji może być utożsamiania z ilością informacji, miary ilości informacji mogą być jednocześnie miarami jej wartości. Oznaczałoby to, że im więcej informacji w ujęciu ilościowym, tym bardziej jest ona wartościowa. Nie zawsze jednak jest to zgodne z prawdą. Załóżmy, że Pani Basia brała udział w loterii, w której prawdopodobieństwo wygrania 100 zł wynosi 0,25. Pan Janusz wybrał natomiast loterię, gdzie szanse na wygranie 100 zł są jak 1:100.

W ujęciu ilościowym: wiadomość o wygranej pani Basi to 2 bity, a o wygranej pana Janusza 6,64 bitów.

W ujęciu wartościowym: wiadomość o wygranej pani Basi ma dla niej dużo większą wartość niż komunikat o wygranej pana Janusza, którego nawet nie zna.

Pieniądz. Pieniądze, szczególnie w przedsiębiorstwach, są częstym miernikiem wartości informacji. Informacja będzie tym bardziej wartościowa, im wyższą kwotę zarobiono lub też zaoszczędzono na skutek jej pozyskania.

Czas. Podobnie jak w przypadku miary, jaką są pieniądze, informacja jest tym bardziej wartościowa, im więcej czasu zaoszczędzono (np. poprzez rezygnację z projektu, użycie nowszych technologii) na skutek jej uzyskania.

Uczestnicy rynku zainteresowani informacją. W tym przypadku im większa liczba uczestników rynku zabiegających o informację, tym bardziej jest ona wartościowa.

Decyzje. O wartości informacji może także świadczyć poprawa skuteczności i trafności podejmowanych decyzji, która nastąpiła na skutek pozyskania informacji. Zatem im większa skuteczność i trafność decyzji, tym bardziej wartościowa jest informacja.

Prawdopodobieństwo osiągnięcia celu. Porównując prawdopodobieństwa osiągnięcia celu po otrzymaniu informacji i przed jej posiadaniem, wskazać można kolejną miarę wartości informacji. Im większa jest bowiem różnica pomiędzy prawdopodobieństwem osiągnięcia celu po otrzymaniu informacji i przed jej odebraniem, tym bardziej informacja ta jest wartościowa (jest to miara powiązana z poprzednią dotyczącą skuteczności i trafności podejmowania decyzji).

Zbliżanie się do celu. Podobnie jak w przypadku prawdopodobieństwa osiągnięcia celu, miarą wartości informacji może być różnica pomiędzy stopniem zbliżenia się do celu po i przed otrzymaniem informacji. Im większa jest ta różnica, tym wyższa jest wartość informacji.

Użyteczność. Dokonując pomiaru i porównania subiektywnej użyteczności rezultatów działań, które zostały podjęte po uzyskaniu informacji i przed jej pozyskaniem, można wskazać kolejną miarę wartości informacji. Informacja jest tym bardziej wartościowa, im większa jest różnica pomiędzy użytecznością rezultatów działań po i przed uzyskaniem informacji[10].

Krotność użycia. Mówiąc o wartości informacji należy pamiętać, że ta sama informacja może być wykorzystywana więcej niż raz. Zatem wartość informacji może być obliczona jako średnia wartość wszystkich możliwych wartości, które wynikają z potencjalnie możliwych sposobów wykorzystania informacji, ważonej przez prawdopodobieństwa zaistnienia każdego takiego faktu.

Podsumowanie

Omówione w pracy sposoby i koncepcje pomiaru ilości i wartości informacji nie wyczerpują tematu. Sygnalizują jedynie problemy związane z tą tematyką i pokazują, jak wiele jeszcze jest do zbadania w teorii informacji.

Literatura

Baborski A., Duda M., Forlicz S., *Elementy cybernetyki ekonomicznej*, PWE, Warszawa 1977.

Baeyer H.Ch. von, *Information. The new language of science*, Harvard University Press, Cambridge, Massachusetts 2004.

Forlicz S., *Informacja w biznesie*, PWE, Warszawa 2008.

Forlicz S., *Mikroekonomiczne aspekty przepływu informacji między podmiotami rynkowymi*, Wydawnictwo Wyższej Szkoły Bankowej w Poznaniu, Poznań 1996.

Forlicz S., *Niedoskonała wiedza podmiotów rynkowych*, Wydawnictwo Naukowe PWN, Warszawa 2001.

Kulikowski J.L., *Informacja i świat w którym żyjemy*, Wiedza Powszechna, Warszawa 1978.

Kuśmierczyk P., *Czy poprawa informacji zawsze prowadzi do poprawy decyzji? Analiza kilku przypadków*, w: *Wpływ niedoskonałej wiedzy podmiotów i ich działalności informacyjnej na funkcjonowanie rynków i gospodarki*, red. S. Forlicz, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 1049, Wrocław 2004.

Mazur M., *Jakościowa teoria informacji*, Wydawnictwa Naukowo-Techniczne, Warszawa 1970.

Nowa encyklopedia PWN, Wydawnictwo Naukowe PWN, Warszawa 1995.

Stefanowicz B., *Informacja*, Wydawnictwo Szkoły Głównej Handlowej, Warszawa 2004.

[10] To samo można powiedzieć o obiektywnych korzyściach.