

Fórum:Modely růstu

Wikiprojekty mají jako mnoho jiných komunitních projektů své životní cykly. Jedná se o období aktivního růstu a období kvantitativního útlumu, který je ovšem vynahrazen zkvalitňováním článků. Některé vlastnosti wikiprojektu se dají matematicky modelovat.

Jednotlivé matematické modely počtu článků

 *Podrobnější informace naleznete na stránce [w:en:Wikipedia:Modelling Wikipedia's growth](https://en.wikipedia.org/wiki/Modelling_Wikipedia's_growth).*

Růst počtu článků ve wikiprojektu má určité zákonitosti, zpočátku roste víceméně exponenciálně, čím více článků ve wikiprojektu je, tím více uživatelů se nadchne a nějaký článek vytvoří. Proto prvním modelem růstu je jednoduchá exponenciála. Toto ovšem netrvá věčně, protože téma, kterým se wikiprojekt zabývá, není nevyčerpatelné, stejně tak někdy nemusí být ani možnosti přispěvatelů. Hovoří se tedy o nějakém limitu, k němuž se pomalu daný parametr růstu pozvolna, asymptoticky blíží. Dva nejpoužívanější modely, které s tímto počítají, jsou Verhulstův model a model Gompertzův.

Verhulstův model

 *Podrobnější informace naleznete na stránce [w:en:Logistic function#In ecology: modeling population growth](https://en.wikipedia.org/wiki/Logistic_function#In_ecology:_modeling_population_growth).*

Verhulstův model, nebo také model logistický, se používá v ekologii a demografii pro modelování růstu populace. Analogických logistických funkcí se používá také např. v neuronových sítích k rozhodnutí, v jakém je neuron stavu.

Velikost, jaké může nějaká populace dosáhnout, je závislá na zdroji potravy, možných příjmech obyvatelstva a dalších parametrech, což je v tomto modelu shrnuto do jediné konstanty **K**. Zároveň populaci charakterizuje ještě další konstanta **r**, která popisuje rychlost růstu. V ekologii je interpretována jako rychlost množení (závisí na délce gravidity atd.), jindy bývá interpretována jako motivace, afinita apod.

V exponenciálním modelu je růst (dP/dt) přímo rychlosti růstu a velikosti populace ($r \cdot P$). Růst ovšem bude ve skutečnosti tím rychlejší, čím bude populace dál od limitu, tedy poměr P/K bude dál od hodnoty jedna. Verhulstův model tedy zavádí korekci:

$$\frac{dP}{dt} = r \cdot P \left(1 - \frac{P}{K}\right)$$

Řešením této diferenciální rovnice a položením integrační konstanty na hodnotu P_0 , která reprezentuje počáteční populaci, získáme základní tvar Verhulstova modelu:

$$P(t) = \frac{KP_0 e^{rt}}{K + P_0(e^{rt} - 1)}$$

Gompertzův model

 *Podrobnější informace naleznete na stránce [w:en:Gompertz function](https://en.wikipedia.org/wiki/Gompertz_function).*

Ukazuje se, že růstu počtu článků ve wikiprojektech ještě o trochu lépe odpovídá Gompertzův model, který se používá mimo jiné např. k modelování růstu nádorové hmoty.

Základním vzorcem Gompertzova modelu je

$$P(t) = a \cdot e^{b \cdot e^{ct}}$$

Konstanta **a** zde udává limit, kapacitu populace, asymptotu, k níž se model blíží, konstanta **b** určuje posun po ose x a konstanta **c** odpovídá rychlostní konstantě **r** z Verhulstova modelu.

Nástroj na výpočet modelů

Stahňte a rozbalte si soubor WikiModely.zip (<https://www.wikiskripta.eu/extensions/javawebstart/verhulstmodel/WikiModely.zip>). K spuštění aplikace musíte mít správně nainstalovanou Javu. Spustíte VerhulstModel.jar a otevřete (Ctrl-O) uložený soubor CSV, který stáhnete na stránce Speciální:GlobalStats. Zvolíte interval parametrů, který by však měl být pro fitování počtu článků již relativně vhodně přednastaven a zmáčnete tlačítko *Run*. Podle rychlosti počítače, velikosti zdrojového souboru dat (CSV) a Vámi předvoleného intervalu parametrů může délka výpočtu trvat od několika vteřin, přes několik dnů a hodin až po věčnost. Přednastavené hodnoty by pro model růstu počtu článků ve WS na běžných PC měly trvat řádově několik minut.

Výstupem programu jsou hodnoty parametrů a graf, který lze uložit.

Metoda výpočtu

Tato aplikace prochází jednotlivé parametry \vec{p} modelu růstu $M(t, \vec{p})$ z předem zvolených intervalů a pro každou kombinaci hodnot spočítá hodnotu *Cost function*, která je definována jako

$$C(\vec{p}) = \sum_{t=0}^n (y_t - M(t, \vec{p}))^2$$

kde y_t jsou skutečná data.

Program potom vybere ty parametry, které mají cost-function nejnižší.

Zdrojový kód aplikace naleznete na GitHubu (<https://github.com/PavelDusek/WikiGrothModels/tree/master/VerhulstModel>).

- WS model - vývoj limitu v čase (<https://docs.google.com/spreadsheets/d/1YF-UbROwSABoVsRPNCuo-0eao8VSI PDIXGTcZHTJdTc/edit#gid=0>)
- editace na článek (<https://docs.google.com/spreadsheets/d/15SEXNbD5izPxziHv9PoLH7bQBkcEVFTAPISwdJ305Do/edit#gid=1197209151>)
- Měření úspěchu wiki (https://docs.google.com/presentation/d/1rTf9isL4l2aAVJiz8bLR6zYKv_W8OffZDoGc_eZPUrw/edit#slide=id.p)
- Medical Educational Wikis (https://docs.google.com/document/d/1UziZ77CG2npGXslt8Ja1z_EjtKdIp54m2iezlhIZdNE/edit)
- podrobnější statistiky wikiprojektů (<http://wikitesty.wz.cz>)