

# SOZIALE BELOHNUNG IN WIKI SYSTEMEN<sup>1</sup>

Bernhard Hoisl, Wolfgang Aigner, Silvia Miksch

*Online-Communities haben eines gemeinsam: Ihr Erfolg steht und fällt mit der Teilnahmerate ihrer Benutzer. Für viele Online-Communities ist der Nutzermangel ein fundamentales Problem. In diesem Artikel wird auf bestimmte Methoden der sozialen Belohnung eingegangen, welche darauf abzielen, Benutzer zu motivieren, in einem Wiki aktiv teilzunehmen. In einer Online-Community basiert soziales Belohnen in den meisten Fällen auf das Hervorheben besonders aktiver Mitglieder. Geld kann dabei nicht als Motivator eingesetzt werden, sondern andere Faktoren wie Status, Macht, Ehre und Ruhm. Es werden verschiedene Methoden der sozialen Belohnung vorgestellt, die die Befriedigung dieser Bedürfnisse als Ziel haben und somit Vorteile für alle Benutzer eines Wikis bringen sollen. Weiters ist eine Anzahl von Methoden sozialer Belohnung in die Wiki Software MediaWiki implementiert worden. Unter anderem sind dabei die Kriterien sozialen Belohnens in der Darstellung einer Rangliste von besonders aktiven Autoren erfüllt. Es wird dabei auf den der Berechnung zugrunde liegenden Algorithmus und den Implementierungsprozess der Entwicklung eingegangen.*

## 1. Einleitung

Die freie Enzyklopädie Wikipedia<sup>2</sup> ist zu einer der größten Wiki Websites gewachsen, in der hunderttausende Benutzer aus allen Teilen der Erde in vielen verschiedenen Sprachen Artikel einstellen und bearbeiten. Diese hohe Teilnehmerrate hat aber zu vielen Problemen, wie etwa der Verbreitung falscher Informationen, Urheberrechtsverletzungen und schlechtem Benutzerverhalten (Spammer, Flamer oder Trolle [17]) geführt. Im Gegensatz zu Wikipedia haben andere Online-Communities große Probleme, ihre Mitglieder zur aktiven Teilnahme zu motivieren. In diesem Artikel werden Techniken vorgestellt mit denen das fundamentale Problem von beiden behandelt wird: Das Erreichen einer kritischen Masse von aktiven Benutzern.

Einerseits hat Wikipedia das Problem, dass veröffentlichte Informationen nicht in einem formalen Prüfungsprozess auf Fehlerfreiheit und Legalität kontrolliert werden. Dadurch bedarf es einer großen und stark involvierten Community, die die bereitgestellte Information freiwillig auf

---

<sup>1</sup> In diesem Artikel wird auf eine geschlechtsneutrale Schreibweise zugunsten der Lesbarkeit verzichtet. Es wird immer der kürzere Ausdruck gewählt, obwohl in *jedem Fall beide* Geschlechter gemeint sind.

<sup>2</sup> <http://www.wikipedia.org>

Korrektheit überprüft. Die Wikipedia-Betreiber haben nicht nur eine soziale, sondern in manchen Fällen auch eine rechtliche Verantwortung, nur fehlerfreie Informationen zu veröffentlichen. Dies auch nicht zuletzt um eine hohe Glaubwürdigkeit von Wikipedia zu ermöglichen. Andererseits haben viele Online-Communities das Problem genug Benutzer zu motivieren um eine lebendige Gemeinschaft aufzubauen. Hohe Mitgliederpartizipation ist ein Schlüsselfaktor für den Betrieb einer erfolgreichen Online-Community. Dies wird in einem Wiki dahingehend noch verstärkt, dass die Produzenten des Gutes (nämlich Information) zugleich auch die Konsumenten sind. Wenn zu wenige Benutzer Inhalte generieren besteht die Gefahr eines Kannibalisierungseffektes. Gute Motivationsfaktoren sind nötig um diesem Phänomen entgegenzuwirken.

Über das Internet bereitgestellte Informationen sind öffentliche Güter. Dadurch können Probleme wie Trittbrettfahrer<sup>3</sup> oder soziales Faulenzen<sup>4</sup> auftreten. Für einen Nutzer eines Wikis erscheint es rational Artikel zu lesen ohne seinerseits einen Beitrag zu leisten. Der Nutzen eines Individuums wird dadurch maximiert, jedoch ist der Gesamtnutzen für die Gemeinschaft nicht optimal. Würde kein Benutzer einen Beitrag leisten (einen Artikel schreiben) wäre der Gesamtnutzen eines Wikis gleich null. Es muss also Autoren geben die zu einem Wiki etwas beisteuern. Wenn ein Autor einen Artikel in Wikipedia einfügt, hat er bestimmte Kosten und will dafür etwas zurückbekommen. Diese Kosten können z.B. Kosten der Informationsbeschaffung, -generierung, -darstellung (vornämlich Zeit), aber auch anfallende Internet Verbindungskosten sein. Den Nutzen eines Autors zu steigern, sodass dieser die auftretenden Kosten übersteigt, ist ein guter Startpunkt um die Partizipationsrate zu erhöhen. Der in diesem Artikel beschriebene Ansatz greift auf Techniken der sozialen Belohnung zurück, um Benutzer zu motivieren, sich aktiv in Online-Communities zu beteiligen.

In dem folgenden Kapitel wird auf verwandte Arbeiten zu diesem Thema eingegangen. Die benutzten Techniken sozialen Belohnens sind in Kapitel 3 beschrieben. In Kapitel 4 wird der Algorithmus der Punkteberechnung, um eine Autorenrangliste zu generieren, erklärt. Kapitel 5 stellt die visuelle Ausgabe der Rangliste vor. Der Implementierungsprozess in MediaWiki wird in Kapitel 6 skizziert und das letzte Kapitel 7 zieht ein Fazit der Arbeit.

## 2. Verwandte Arbeiten

Es sind zahlreiche Artikel und Bücher über das Phänomen Wiki publiziert worden (vergleiche z.B. [1, 6, 9, 18]). Die meisten Arbeiten beziehen sich jedoch auf das Installieren und Betreiben eines Wikis oder dem Versionsverwaltungssystem und dessen Vorteile für gemeinschaftliches Arbeiten. Leider wird der Untersuchung des Benutzerverhaltens in Online-Communities sehr wenig Aufmerksamkeit gewidmet. Einige Forschungsansätze beschäftigen sich mit dem Problem des Trittbrettfahrens im Internet [1, 5]. Es existieren auch Studien über das Kommunikationsverhalten

---

<sup>3</sup> Als Trittbrettfahrer wird ein Individuum bezeichnet, das den Nutzen eines Gutes erlangt, ohne dafür zu bezahlen. Trittbrettfahrerverhalten kann nur bei Gütern auftreten, bei denen keine Ausschließbarkeit möglich ist, z.B. bei reinen öffentlichen Gütern [15]. Es gäbe keine Trittbrettfahrer wenn jeder Nutzer denselben Teil an Informationen zu einem Wiki beiträgt. Eines der größten Probleme dabei ist, dass der Wert einer Information für ein Individuum sehr subjektiv und nur schwer feststellbar ist.

<sup>4</sup> „Soziales Faulenzen [...] bedeutet, dass Menschen in Gruppen [...] weniger leisten, als wenn sie die Aufgabe allein bewältigen“ [11]. Das Problem des sozialen Faulenzens tritt deswegen so häufig auf, weil die wenigsten Artikel in Wikipedia (wie in den meisten anderen Wikis) von einer Person allein geschrieben werden. Die Antwort auf soziales Faulenzen sind motivierende Faktoren. Ein erster (nicht so beabsichtigter) Ansatz in MediaWiki ist, dass jeder Teil eines Artikels mit genau einem Autor verknüpft ist und das Geschriebene somit bewertet werden kann.

von Nutzern in Online-Communities [18], aber der Fokus wird dabei nicht auf für Benutzer motivierende Faktoren gelegt.

Abraham Maslow diskutierte bereits sehr früh die für einen Menschen motivierenden Faktoren in seiner Bedürfnispyramide [9]. Darauf und auf sozialpsychologische Einflüsse basieren zwei Artikel die sich mit den Problemen von geringer Benutzerpartizipation und dem sozialen Faulenzen in einer Online-Community beschäftigen [6, 8].

Wie bereits erwähnt zielt unser Ansatz der Benutzermotivation auf das Hervorheben besonders aktiver Autoren und deren Reputation ab [16]. Wir wollen Mitglieder einer Online-Community motivieren, damit diese die selbstständig eingegebenen Informationen überprüfen und gegebenenfalls korrigieren. Dass so ein Ansatz einer mitgliedergesteuerten Community die Quantität und Qualität der Beiträge erhöht, wurde in anderen Arbeiten bestätigt [3] und anhand von Wikipedia empirisch getestet [2].

### 3. Techniken sozialer Belohnung

In diesem Artikel werden Techniken sozialen Belohnens präsentiert, welche einen Mehrwert für die Mitglieder einer Online-Community bringen sollen, sodass eine höhere Partizipationsrate erzielt werden kann. Soziale Belohnung („*social rewarding*“) bezieht sich dabei auf die Erhöhung der Intensität eines bestimmten Verhaltens. In Online-Communities basiert soziales Belohnen in den meisten Fällen auf das Hervorheben besonders aktiver Mitglieder. Es sind vier Methoden sozialen Belohnens in die Software MediaWiki<sup>5</sup> (welche auch von Wikipedia benutzt wird) implementiert worden. Durch den Einsatz der ersten drei Mechanismen kann eine Rangliste der aktivsten Autoren erstellt werden. Die vierte Methode ist ein Empfehlungssystem, um Autoren und Benutzer mit denselben Interessen vorzuschlagen. Dabei beziehen sich die hier vorgestellten Techniken hauptsächlich auf die automatische Untersuchung von quantitativen und qualitativen Charakteristika publizierter Artikel:

- *Anzahl an Referenzen* („*Amount of References*“): Mittels Google’s SOAP Search API<sup>6</sup> wird ein Artikel auf folgende drei Kriterien untersucht: Größe benutzter Referenzen, Anzahl an Verweisen, die auf die benutzten Referenzen zeigen und die Anzahl an eingehenden Verweisen auf den Wiki-Artikel.
- *Bewertung von Artikeln* („*Rating of Articles*“): Für Benutzer fehlt in MediaWiki die Möglichkeit einen Artikel bewerten zu können. Wir implementierten ein offenes Bewertungssystem, mit dem Nutzer für oder gegen einen Artikel stimmen und optional auch einen Kommentar hinterlassen können.
- *Meist angesehene Artikel* („*Most Viewed Articles*“): Artikelbesuche von Benutzern werden, basierend auf vordefinierte Kriterien, gezählt.
- *Empfehlungen* („*Recommender System*“): Aufgrund des Benutzerverhaltens (welche Artikel angesehen wurden) werden Empfehlungen abgegeben. Es gibt drei Arten von Vorschlägen: empfohlene Artikel oder Artikelversionen<sup>7</sup>, empfohlene Autoren und Benutzer mit denselben Interessen.

---

<sup>5</sup> <http://www.mediawiki.org>

<sup>6</sup> Vereinfacht dargestellt ist die SOAP Search API von Google eine definierte Schnittstelle mit der automatisierte Suchabfragen in den eigenen Programmcode eingebunden werden können.

<sup>7</sup> Jede Änderung, die an einem Artikel vorgenommen wird, erzeugt eine neue Artikelversion („*revision*“).

Die zwei wichtigsten Kriterien für die Auswahl gerade dieser drei Techniken war einerseits eine gute Kombination verschiedener Methoden zu finden und andererseits die Reduktion der Komplexitätsstufe des Implementierungsprozesses in MediaWiki. Unsere Entwicklung wurde als eine Erweiterung („*extension*“) zu MediaWiki konzipiert. Somit kann jeder Betreiber eines MediaWiki Systems diese ohne große Probleme in sein eigenes Wiki integrieren. Es gibt zwei Gründe warum wir glauben, dass der Einsatz mehrerer verschiedener Techniken sozialer Belohnung die besten Resultate liefert. Der erste Grund ist, dass beim Einsatz verschiedener Methoden Daten aus verschiedenen Quellen gesammelt werden. Die Kombination dieser Daten sollte bessere Ergebnisse liefern als die Nutzung nur einer Technik allein. Der zweite Grund ist, dass es die Benutzung mehrerer Datenquellen für Autoren schwerer macht, zu betrügen. Wenn nur die Artikelaufrufe von Benutzern gezählt werden, dann ist es möglich, dass Autoren ihre Artikel oft selber besuchen (möglicherweise unter einem anderen Login) und damit das Ergebnis verfälschen. Natürlich müssen Kontrollmechanismen solche Handlungsweisen bis zu einem gewissen Grad verhindern bzw. erschweren, jedoch gehen solche Einschränkungen immer zu Lasten des *wirklichen* Verhaltens der Benutzer, welches wir zu messen versuchen.

### **3.1. Anzahl an Referenzen**

Wir gehen davon aus, dass der Wert eines Artikels mit der Anzahl und Reputation der verwendeten Referenzen wächst. Ein Ansatz eines automatisierten Qualitätschecks wurde mit Google<sup>8</sup> (einer der weltweit größten Suchmaschinen) realisiert. Wie bereits angesprochen, sind dafür drei Kriterien maßgebend. Es soll herausgefunden werden wie viele Seiten auf die in dem Artikel benutzten Referenzen zeigen<sup>9</sup>. Wenn sehr viele Seiten mit einer Referenz verlinkt sind und diese Seiten wiederum viele Links auf sich ziehen, dann scheint die Ressource eher vertrauenswürdig, als wenn nur wenige oder gar keine Verlinkungen stattgefunden haben. Auf jeden Fall hat diese Referenz ein gewisses Maß an Glaubwürdigkeit. Weiters wird auch die Größe der Referenz in die Berechnung miteinbezogen<sup>10</sup>. Wir gehen davon aus, dass eine Referenz mit tausenden von Unterseiten vertrauenswürdiger ist, als eine kleine private Website mit nur wenigen Seiten. Als letztes Kriterium wird die Anzahl der Links von außerhalb des Wikis, die auf den untersuchten Artikel zeigen, gezählt. Wir nehmen an, dass ein Artikel mit vielen Verlinkungen einen hohen Wert für viele Leute hat. Je höher die Anzahl der Verweise, desto höher ist auch die Zahl der Benutzer, die sich diesen Artikel ansehen. Viele Links können also ein Indikator für einen gut geschriebenen Artikel sein. Mit dieser Methode kann die Bekanntheit eines Artikels gemessen und ein erster Qualitätscheck erfolgen.

### **3.2. Bewertung von Artikeln**

Um gute von schlechten Artikeln zu unterscheiden, hat der Benutzer die Möglichkeit dafür oder dagegen zu stimmen. Dazu wird eine Punkteskala definiert, aufgrund dessen der Benutzer abstimmt. Weiters kann auch ein optionaler Kommentar abgegeben werden, welcher auf der Diskussionsseite des jeweiligen Artikels erscheint. Der Autor hat nun die Möglichkeit zu sehen, warum Benutzer ihn gut oder schlecht bewertet haben und kann aufgrund dieser Erkenntnis den Artikel umschreiben (natürlich kann jeder andere Benutzer dies auch selber tun). Für den Autor ist auf der Diskussionsseite außerdem Platz, um auf Benutzerkommentare zu reagieren und sich gegebenenfalls zu rechtfertigen. Somit werden die Punkte quantitativ gezählt, während der Autor

---

<sup>8</sup> <http://www.google.com>

<sup>9</sup> Die Eingabe von `link:http://www.tuwien.ac.at` als Suchbegriff gibt alle Seiten zurück die auf diese Adresse zeigen.

<sup>10</sup> Durch eine Suchanfrage mit dem Begriff `site:http://www.tuwien.ac.at` kann die Anzahl der Unterseiten der Adresse herausgefunden werden.

mit Hilfe des Kommentars auch eine qualitative persönliche Rückmeldung erhält. Das Aussehen der Bewertung und dessen Platzierung bleibt dem Autor des Artikels überlassen, wobei es einen konfigurierbaren Standard gibt. Ein weiterer zu berücksichtigender Aspekt ist die Anzahl an Bewertungen für ein repräsentatives Ergebnis und die Verhinderung von Mehrfachbewertungen.

### **3.3. Meist angesehene Artikel**

Die Idee hinter einer Liste von meist angesehenen Artikeln ist, dass, wenn ein Artikel von vielen Leuten betrachtet wird, er entweder (1) sehr informativ und gut geschrieben ist, oder ein (2) hochinteressantes Thema für eine große Masse an Personen beinhaltet. Im ersten Fall kann gesagt werden, dass ein Artikel mit hohen Besuchszahlen einem Autor hilft, eine gute Reputation zu erlangen. Meist angesehene Artikel können nach Zeit oder Thema aufgeschlüsselt werden. Bei den meist angesehenen Artikeln wird es für die gesamte Lebensdauer eines Wikis nicht viele Schwankungen geben, weshalb eine Unterteilung in meist angesehene Artikel der Woche oder des Monats eine gute Lösung sein können.

### **3.4. Empfehlungen**

Empfehlungssysteme versuchen Objekte vorauszuberechnen, die ein Benutzer interessant finden könnte. Die Berechnung basiert dabei auf Daten des bisherigen Benutzerverhaltens und -profils. Ein Empfehlungssystem gibt Vorschläge für einen einzelnen Benutzer ab, welche auf Informationen basieren, die von vielen Benutzern gesammelt wurden. In unserer konkreten Implementierung basieren die Empfehlungen auf einer einstellbaren Anzahl von Top-Artikeln eines Benutzers. Top-Artikel eines Benutzers sind jene Artikel, die ein Benutzer am häufigsten besucht hat. Für jeden Artikel in der Top-Artikelliste eines Autors werden Übereinstimmungen mit demselben Artikel in der Top-Artikelliste aller anderen Autoren gesucht. Somit ergibt sich eine Liste mit Artikeln, die nach den Besuchen des gesuchten Artikels und nach den Besuchen der Artikel in der Top-Artikelliste des gefundenen Autors gewichtet werden. Wird diese gewichtete Liste nun absteigend sortiert, dann sind die empfohlenen Artikel, Artikelversionen oder Autoren an oberster Stelle. Es kann konfiguriert werden wie viele der empfohlenen Objekte angezeigt werden sollen. Um eine Aufstellung mit Benutzern mit denselben Interessen zu bekommen, muss die erstellte Liste mit empfohlenen Objekten eines Benutzers mit den Listen der anderen Benutzer verglichen werden um Übereinstimmungen zu finden. Benutzer mit denselben Interessen sind die Benutzer mit den meisten Übereinstimmungen, wobei die Kriterien der Übereinstimmung wiederum einiger Gewichtungen unterzogen werden. Zum Schluss muss erneut definiert werden wie viele Objekte angezeigt werden sollen.

Durch ein Empfehlungssystem können also Autoren für einen Benutzer angezeigt werden, die für diesen Benutzer interessant erscheinen. Wenn ein Benutzer zu einem bestimmten Thema einen Artikel sucht, kann er mit dieser Technik alle anderen Autoren, die sich mit gleichen oder ähnlichen Dingen beschäftigen kennenlernen. Somit wird bei dieser Methode das Hervorheben von Autoren nicht aufgrund der Partizipationsrate, sondern durch den Vergleich von Ähnlichkeitskriterien durchgeführt.

Wie nun die ersten drei Techniken sozialer Belohnung (*Anzahl an Referenzen, Bewertung von Artikeln, Meist angesehene Artikel*) zusammenhängen, um eine Autorenrangliste zu generieren, zeigt das nächste Kapitel.

## 4. Punkteberechnung

Die Berechnung einer Liste der aktivsten Autoren erfolgt in zwei Stufen. Zuerst werden auf Basis der Versionen eines Artikels Punkte berechnet und danach den Autoren zugeteilt.

### 4.1. Berechnung aufgrund der Artikelversion

Jede der drei Methoden sozialen Belohnens berechnet Punkte für jede Artikelversion in einem Wiki. Dies wird erreicht, indem der Wert der untersuchten Version mit dem Durchschnittswert aller Artikelversionen in Beziehung gesetzt wird (Formel 1).

$$avg_{R_j} = \frac{\sum_{i=1}^n r_{ij}}{n} \quad (1)$$

$R$  ist dabei eine Menge von allen Versionen, wobei  $r_{ij}$  der Wert der  $j$ -ten Technik sozialer Belohnung der Version  $i$  ist:

$$R_i = \{r_{i1}, r_{i2}, r_{i3}, \dots, r_{ij}; 1 \leq j \leq 3\}. \quad (2)$$

Für z.B. die Methode *Meist angesehene Artikel* werden alle Besuche einer Artikelversion gezählt. Wir nehmen an, dass Artikel *C* in der 6. Version 20 Besucher hat. Wenn nun der Durchschnitt der Besucher aller Artikelversionen 30 wäre, dann hat Artikel *C* 66.67% Besucher. Wir wollen jeder Version aufgrund ihrer Besucherzahlen Punkte zuteilen, weshalb eine Punkteskala definiert werden muss. In diesem Beispiel würden 66.67% Besucher 2 Punkte von 5 möglichen bedeuten<sup>11</sup>.

Dieses Beispiel der Punktezurechnung wird für jede Methode und für jede Artikelversion vorgenommen. Bei der Technik *Bewertung von Artikeln* bewerten Benutzer die Versionen der Artikel von 0 bis 5 Punkten. Für die Methode *Anzahl an Referenzen* werden die Größe benutzter Referenzen, die Anzahl an Verweisen die auf die benutzten Referenzen zeigen und die Anzahl an Verweisen auf den Artikel selbst für die Berechnung herangezogen. Diese drei Kriterien können nach einer benutzerdefinierten Skala gewichtet werden, um dann wiederum den Vergleich mit einem Durchschnittswert anzutreten.

Am Ende des ersten Berechnungsschrittes gibt es für jede Artikelversion und jede Technik sozialer Belohnung einen Punktwert aufgrund der vordefinierten Skala. Diese Werte werden gewichtet und auf einen Gesamtwert pro Artikelversion aufsummiert. Bei der Betrachtung von Formel 3 kann man sehen, dass  $p_{r_i}$  die aufsummierten Punkte für Artikelversion  $i$  jeder Technik sozialer Belohnung  $j$  gewichtet nach  $w_j$  (welches in der Konfigurationsdatei definiert werden muss) ist.

$$p_{r_i} = \sum_{j=1}^3 \frac{r_{ij}}{avg_{R_j}} * w_j \quad (3)$$

Die Punktezuweisung zu Autoren erfolgt in der nächsten Stufe.

---

<sup>11</sup> In den gezeigten Beispielen wird eine Skala von 0 (am schlechtesten) bis 5 (am besten) Punkten benutzt. Diese kann jedoch frei konfiguriert werden.

## 4.2. Punktezuteilung zu Autoren

Eine Artikelversion hat genau einen Autor, weshalb eine Summierung aller Punkte aller geschriebenen Artikelversionen eines Autors möglich wird. Die Punkte werden dazu nach zwei unterschiedlichen Methoden gewichtet: Nach der Änderungslänge der Artikelversion und nach dem Änderungsdatum. Dazu wird zuerst eine veränderte Menge von  $R$  definiert, wobei  $r'_{ik}$  die Artikelversion  $k$  des Artikels  $i$  ist (Formel 4).

$$R'_i = \{r'_{i1}, r'_{i2}, r'_{i3}, \dots, r'_{ik}\} \quad (4)$$

Wir gehen davon aus, dass die Größe einer Änderung in einer Artikelversion im Vergleich zur vorhergehenden proportional zum Wert dieser Änderungen ist. Dabei spielt es keine Rolle ob die neue Version erweitert oder gekürzt wurde – ein Mehrwert der Artikelqualität wird angenommen<sup>12</sup>). Die Änderung einer Version zur vorhergehenden wird in Bytes gezählt. Formel 5 berechnet einen Gesamtwert aller Artikeländerungen aller Artikelversionen  $k$  von einen Artikel  $i$  (wobei  $s_{r'_{ik}}$  die Artikeländerung einer Artikelversion im Vergleich zur vorhergehenden ist).

$$S_{r'_i} = \sum_{k=1}^n \underbrace{abs[size(r'_{ik}) - size(r'_{i(k-1)})]}_{s_{r'_{ik}}} \quad (5)$$

Die zweite Annahme ist, dass neuere Artikelversionen mehr zählen als ältere. Wir gehen davon aus, dass neuere Artikelversionen aktuellere Themen behandeln und deshalb eine höhere Gewichtung haben sollen, als solche die vor vielleicht zwei oder drei Jahren geschrieben wurden. Formel 6 summiert dafür die relative Zeitspanne aller Artikelversionen  $k$  für einen Artikel  $i$  (wobei  $t_{r'_{ik}}$  das relative Änderungsdatum einer Artikelversion ist).

$$T_{r'_i} = \sum_{k=1}^n \underbrace{time(r'_{ik}) - firstTime(r'_i)}_{t_{r'_{ik}}} \quad (6)$$

Somit werden für jede Version jedes Artikels die Unterschiede zur vorhergehenden Version und das Alter verglichen mit dem Erstellungsdatum des Artikels gespeichert (Formeln 5 und 6).

Nun wird eine Untermenge  $A$  definiert (Formel 7) welche alle Artikelversionen eines Autors beinhaltet. Folglich werden nur Artikelversionen von dem Autor, für den die Berechnung durchgeführt wird, benutzt. D.h., dass z.B.  $s_{a_{ik}}$  (in Formel 8) die Artikeländerung einer Artikelversion eines Autors ist, welche durch die Gesamtänderung aller Artikelversionen des Artikels dividiert wird, um einen Prozentsatz zu errechnen.

$$A \subseteq R' \quad (7)$$

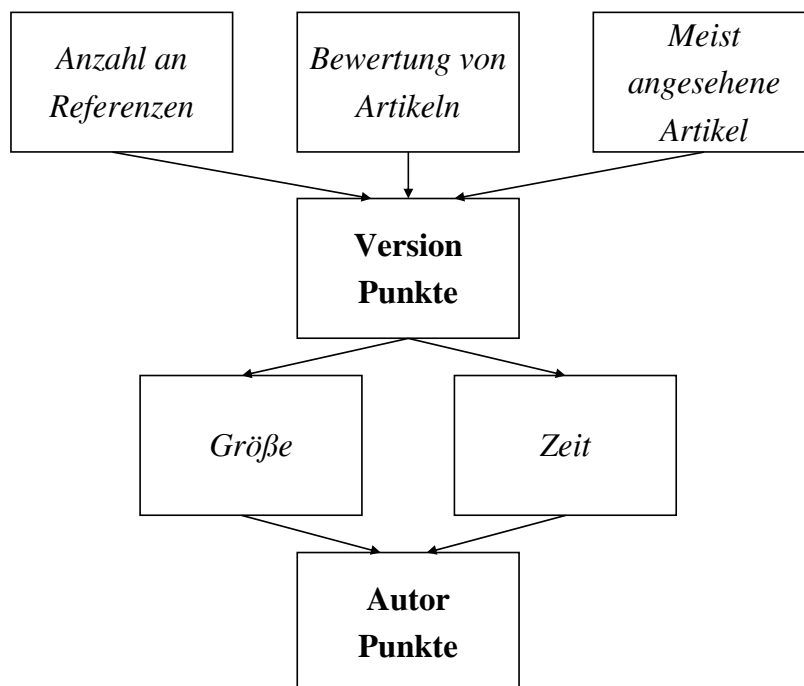
In Formel 8 werden für jede Artikelversion, die zu einem Autor gehört und für jedes Kriterium (Artikeländerung und Zeit) Prozentzahlen errechnet, welche anhand einer vordefinierten Skala gewichtet werden ( $w_S$  and  $w_T$ ) – je nachdem wie wichtig ein Faktor für die Punkteberechnung sein soll. Dann werden die beiden Werte mit den im ersten Schritt für diese Artikelversion berechneten

<sup>12</sup> D.h. das Verhalten von Spammern, Flamern, Trollen etc. wird nicht berücksichtigt.

Punkten multipliziert ( $p_{r_k}$ ) und summiert. Das Ergebnis ist eine gewichtete Punkteanzahl für jede Artikelversion ( $p_{a_{ik}}$ ), welche für jeden Artikel, der von einem Autor geschrieben wurde, zusammengezählt werden muss ( $p_A$ ).

$$p_A = \sum_{i=1}^n \sum_{k=1}^m \underbrace{\frac{s_{a_{ik}}}{S_{r_i}} * w_S * p_{r_k} + \frac{t_{a_{ik}}}{T_{r_i}} * w_T * p_{r_k}}_{p_{a_{ik}}} \quad (8)$$

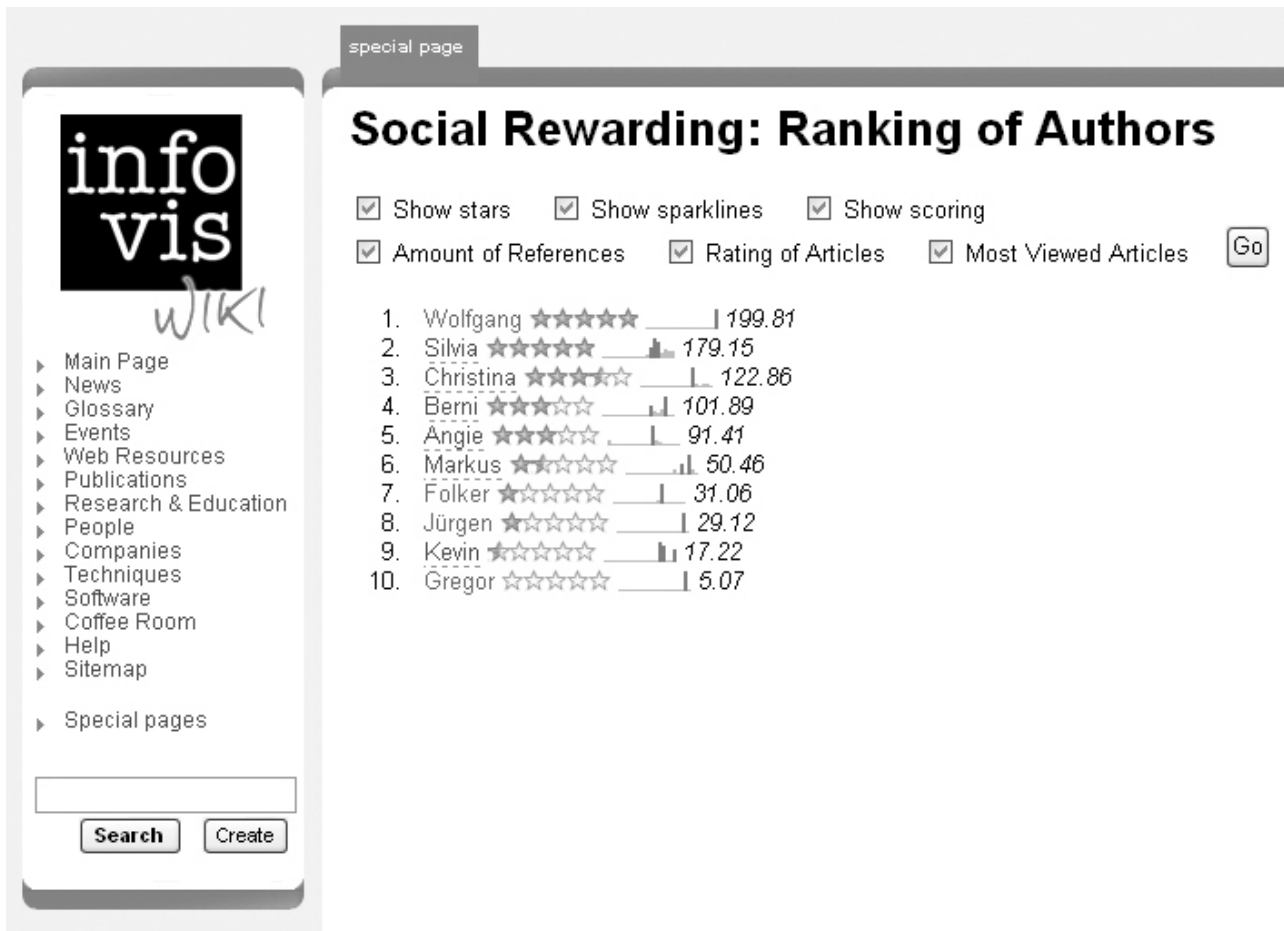
Diese Berechnung erfolgt für jeden Autor separat, sodass am Ende jeder Autor einen Punktwert zugeteilt bekommt, welcher als Basis für die Autorenrangliste dient. Abbildung 1 gibt nochmals einen Überblick über die zweistufige Punkteberechnung.



**Abbildung 1:** Schematische Darstellung der zweistufigen Punkteberechnung. Zuerst werden Punkte auf Basis der Artikelversionen für alle drei Techniken sozialer Belohnung berechnet. In der zweiten Stufe werden diese Punkte nach Größe der Änderung und zeitlichen Kriterien gewichtet und für alle Autoren summiert.

## 5. Autorenrangliste

Um die Ergebnisse der Berechnung darzustellen, können verschiedene Autorenranglisten erzeugt werden, bei denen der aktivste Autor an erster Stelle stehen wird. Um die graphische Darstellung zu unterstützen werden zwei bekannte Techniken der Informationsvisualisierung eingesetzt: *Sterne* und *Sparklines* [20] (Abbildung 2).



**Abbildung 2: Screenshot einer Autorenrangliste. Neben den Autorennamen sind Sterne, Sparklines und die errechneten Punkte zu sehen.**

### 5.1. Sterne

Sterne werden gerne als graphische Darstellung eingesetzt um schnell zu sehen, wie gut oder schlecht etwas ist. Große Internetseiten wie eBay oder Amazon und viele Foren benutzen Sterne als Darstellungsform. Wir verwenden dazu die am weitesten verbreitete fünfstufige Skala, wobei die Anzeige halber Sterne zusätzlich aktiviert werden kann. Da Sterne auf Basis der Partizipationsrate aller Mitglieder der Gemeinschaft berechnet werden, sind sie ein guter Indikator für die Höhe der Gesamtteilnahme eines Benutzers.

### 5.2. Sparklines

Sparklines sind kleine, hochauflösende Grafiken, die leicht in einem fortlaufenden Text integriert werden können. Einsatzmöglichkeiten sind z.B. bei der Darstellung von Finanz- und Wirtschaftsdaten oder bei der Veranschaulichung von Veränderungen über einen gewissen Zeitraum gegeben [20]. In unserem Fall werden Sparklines für die Anzeige der Partizipationsrate eines Benutzers über eine bestimmte Zeitspanne, unterteilt in vordefinierte Intervalle, benutzt. Sparklines wurden als Darstellungsmittel vornämlich wegen ihrer guten Integration in Fließtexten und ihrer Schlichtheit gewählt. Dunklere Sparklines zeigen eine überdurchschnittliche Partizipationsrate, hellere eine unterdurchschnittliche an. Das Aussehen, die Intervalle, die Farben etc. können vom Benutzer angepasst werden.

## 6. Implementierung

### 6.1. MediaWiki

Die in den vorigen Kapiteln beschriebenen Techniken sozialer Belohnung wurden in MediaWiki als Erweiterung implementiert. MediaWiki ist eine open-source Wiki Software welche unter der GNU General Public License (kurz GPL) vertrieben wird. Basis ist die Programmiersprache PHP und MySQL wird als Datenbankmanagementsystem eingesetzt. Wir benutzen MediaWiki als Kernsystem wegen der hohen Stabilität und guten Fähigkeit Software von Drittanbietern zu integrieren.

#### 6.1.1. Erweiterung

Die implementierte Erweiterung in MediaWiki hat den Namen *SocialRewarding*. Jeder mit einer funktionierenden MediaWiki Installation kann diese Erweiterung in sein Wiki integrieren. Die Installation der Erweiterung wurde dabei so einfach wie möglich gehalten und erfordert kein tiefgreifendes Wissen der beteiligten Systeme.

#### 6.1.2. Special Pages

Special Pages sind Wiki-Seiten, die nicht in die reguläre MediaWiki-Struktur passen (keine Artikel oder Diskussionsseiten). Einige Special Pages sind von MediaWiki bereits vorinstalliert und stellen Funktionen wie z.B. Benutzerlisten oder zuletzt bearbeitete Artikel bereit. Special Pages benutzen einen eigenen Namespace *Special:* und sind nicht so wie andere Seiten direkt editierbar [12]. Alle von der SocialRewarding Erweiterung erstellten Seiten sind Special Pages (beginnend mit dem Präfix „*SocialRewarding*“) und über die Liste aller Special Pages erreichbar.

#### 6.1.3. Hooks

Wenn eine Funktion einem Hook zugeteilt wird, dann wird diese Methode an einer definierten Stelle des MediaWiki-Codes aufgerufen, um die in der Funktion beschriebenen Routinen auszuführen. Damit bekommt der Entwickler die Möglichkeit eigene Methoden in die Laufzeitausführung von MediaWiki zu integrieren [13]. Die SocialRewarding-Erweiterung benützt Hooks, z.B. wenn ein Artikel gespeichert wird, um eine Funktion aufzurufen, die die zitierten Referenzen des Artikels sucht und diese mittel Google überprüft (so wie in vorigen Kapiteln beschrieben).

#### 6.1.4. Markups

Wiki-Markup ist die Syntax, die ein Autor benutzen kann um einen Artikel zu formatieren [14]. Die SocialRewarding-Erweiterung erstellt einige Markups, welche vordefinierten Text in einen Artikel lädt, z.B. ein Bewertungsformular, mit dessen Hilfe ein Benutzer dem Artikel Punkte vergeben kann. Für diesen Zweck werden XML-Tags definiert, die im Wiki-Editor wie folgt benutzt werden:  
<tagname> irgendein Text </tagname>

## 6.2. SocialRewarding-Erweiterung

### 6.2.1. Objekte

Die SocialRewarding-Erweiterung hat eine objektorientierte Struktur:

- Klasse *ManageReward*: Verwaltet die Methoden zur Datenzwischenspeicherung („*cache*“).
- Klasse *Reward*: Das ist die Hauptklasse zur Berechnung der Punkte der ersten drei Techniken sozialer Belohnung und zur Anzeige der Autorenrangliste mit Sternen und Sparklines.
- Klassen *RewardReferences*, *RewardRating*, *RewardViewed* und *RewardRecommend*: Jede Methode sozialen Belohnens wird durch eine eigene Klasse repräsentiert, welche wiederum Unterklassen besitzen können, z.B. *GoogleSearch* für die Verwaltung der Suchanfragen.

### 6.2.2. Zwischenspeicher

Da die Berechnung der Autorenrangliste und der Autorenempfehlungen sehr stark von der Anzahl der vorhandenen Artikel, Artikelversionen und Autoren in einem Wiki abhängt, kann diese sehr zeitintensiv werden. Ein Zwischenspeicher wurde entwickelt um die Berechnung nicht bei jeder Anfrage ausführen zu müssen. Da alle nötigen Daten zur Berechnung der Autorenrangliste und der Autorenempfehlungen in Klassen gehalten werden, kann eine speicherbare Repräsentation dieser Objekte erstellt werden. Der Speicherort ist konfigurierbar, entweder auf der Festplatte oder in der Datenbank. Wenn als Zwischenspeicher die Datenbank gewählt wird, ist es möglich eine Historie von Autorenranglisten zu erstellen. Dabei ist zu beachten, dass mitunter sehr viele Daten in die Datenbank geschrieben werden und deshalb vermutlich der Kommunikationspuffer der Datenbank erhöht werden muss.

### 6.2.3. Datenbank

Die Integration der SocialRewarding-Erweiterung betrifft nicht alle Tabellen der MediaWiki-Datenbank. Abhängigkeiten existieren nur mit vier Tabellen: *revision*, *page*, *user* und *text*. Für die Speicherung benötigter Daten erstellt die SocialRewarding-Erweiterung bei der Installation sechs neue Tabellen.

### 6.2.4. Installation

Die SocialRewarding-Erweiterung muss geladen werden, indem die MediaWiki-Datei `LocalSettings.php` editiert wird. Eine Special Page führt die Installation durch, in der auch eine erste Dateninitialisierung stattfinden kann.

### 6.2.5. Dateninitialisierung

Eine Initialisierung bereits existierender Daten ist nötig, weil die SocialRewarding-Erweiterung am Anfang keinerlei bestehende (Vergleichs)daten hat und daher einige Zeit benötigt um gute Resultate zu liefern. Deswegen werden vorhandene Daten von MediaWiki transformiert, um eine erste Berechnung durchführen zu können. Daten werden nur durch die aktive Teilnahme von Mitgliedern des Wikis gesammelt. Deswegen kann es Tage oder sogar Wochen dauern (abhängig von der Partizipationsrate der Benutzer) bis genug Daten gesammelt wurden, um repräsentative Ergebnisse zu liefern. Bevor die Erweiterung installiert und Daten initialisiert werden, ist es ratsam einen Blick auf die Konfigurationsdatei zu werfen.

### 6.2.6. Konfiguration

Es existieren zwei Konfigurationsdateien, die sich in dem Detaillierungsgrad der einzelnen Funktionsbeschreibungen unterscheiden. Ungefähr 100 verschiedene Variablen können gesetzt werden, um den Benutzerbedürfnissen der Online-Community Rechnung zu tragen.

## 7. Fazit

Nutzermangel ist ein ernstes Problem für viele Online-Communities. In diesem Artikel wurde versucht einen Einblick zu geben, wie es gelingen kann, durch Techniken sozialer Belohnung Benutzer zu motivieren. Unsere Arbeit basierte dabei auf das Hervorheben besonders aktiver Mitglieder einer Wiki-Gemeinschaft und im Fall des Empfehlungssystems auf das Anzeigen vorgeschlagener Autoren. Um die aktivsten Benutzer zu finden, wurde eine Autorenrangliste mit Hilfe von drei entwickelten Methoden sozialen Belohnens erstellt: *Anzahl an Referenzen*, *Bewertung von Artikeln* und *Meist angesehen Artikel*. Zahlreiche Variablen beeinflussen die Berechnung, wobei einige davon konfigurierbar sind, andere von der Qualität, Quantität und Aktualität eines Artikels abhängen. Neben der Autorenrangliste werden Sterne und Sparklines zur besseren Visualisierung der Ergebnisse benutzt. Implementiert wurden die Techniken in eine open-source Wiki-Software namens MediaWiki.

Unser Ansatz kann als Startpunkt zur Entwicklung von Mechanismen für die Motivation von Nutzern einer Online-Community angesehen werden. Wir glauben, dass das Zusammenwirken von mehreren Techniken sozialer Belohnung ein guter Weg ist, um qualitativ hochwertige Ergebnisse zu liefern, welche benötigt werden, um ausreichend Anreize für Benutzer zu schaffen. Trotzdem ist es ein Fehler anzunehmen, dass ausschließlich die hier behandelten Methoden ausreichen, um genug Benutzer zu motivieren und eine aktive Gemeinschaft für jedes erdenkliche Wiki zu schaffen. Nutzer müssen vor allem intrinsisch motiviert sein um sich an einem Wiki zu beteiligen, was durch unsere Techniken nur bis zu einem gewissen Grad angeregt werden kann.

Das hier vorgestellte Projekt ist noch nicht öffentlich zugänglich gemacht worden, weshalb auch noch keine empirischen Daten vorhanden sind. Es ist jedoch von uns angedacht, dass unsere Entwicklung über das Internet kostenlos zu beziehen sein wird. Zurzeit wird daran gearbeitet, die beschriebenen Konzepte in einem größeren Maßstab einer Evaluierung zu unterziehen.

## 8. Literatur

- [1] Adar, Eytan; Huberman, Bernardo A.: Free Riding on Gnutella. Xerox Palo Alto Research Center (2000).
- [2] Chesney, Thomas: An Empirical Examination of Wikipedia's Credibility. First Monday, volume 11, number 11 (2006).
- [3] Cosley, Dan; Frankowski, Dan; Kiesler, Sara; et. al.: How Oversight Improves Member-Maintained Communities. CommunityLab (2005).
- [4] Ebersbach, Anja; Glaser, Markus; Heigl, Richard: Wiki: Web Collaboration. Springer, Berlin Heidelberg New York (2005).
- [5] Feldman, Michal; Papadimitriou, Christos; Chuang, John; et. al.: Free-Riding and Whitewashing in Peer-to-Peer Systems. School of Information Management and Systems, UC Berkeley (2004).

- [6] Kimberly, Ling; Beenen, Gerard; Ludford, Pamela; et. al.: Using Social Psychology to Motivate Contributions to Online Communities. CommunityLab (2004).
- [7] Leuf, Bo; Cunningham, Ward: The Wiki Way: Quick Collaboration on the Web. Addison-Wesley Professional (2001).
- [8] Ludford, Pamela; Cosley, Dan; Frankowski, Dan; et. al.: Think Different: Increasing Online Community Participation Using Uniqueness and Group Dissimilarity. University of Minnesota, Department of Computer Science and Engineering (2004).
- [9] Maslow, Abraham H.: Motivation and Personality. 3<sup>rd</sup> ed. HarperCollins Publishers (1987).
- [10] McFederation, Paul: Technically Speaking: It's A Wiki, Wiki World. In: IEEE Spectrum, volume 43, 88 (2006).
- [11] o.A.: Gruppenleistung. <http://de.wikipedia.org/wiki/Gruppenleistung>, abgefragt am 13. Februar 2007.
- [12] o.A.: Help: Special Pages. [http://www.mediawiki.org/wiki/Help:Special\\_pages](http://www.mediawiki.org/wiki/Help:Special_pages), abgefragt am 13. Februar 2007.
- [13] o.A.: Manual: MediaWiki Hooks. [http://www.mediawiki.org/wiki/Manual:MediaWiki\\_hooks](http://www.mediawiki.org/wiki/Manual:MediaWiki_hooks), abgefragt am 13. Februar 2007.
- [14] o.A.: Wikipedia: How to Edit a Page. [http://en.wikipedia.org/wiki/How\\_to\\_edit\\_a\\_page](http://en.wikipedia.org/wiki/How_to_edit_a_page), abgefragt am 13. Februar 2007.
- [15] o.A.: Trittbrettfahrerverhalten. <http://de.wikipedia.org/wiki/Trittbrettfahrerverhalten>, abgefragt am 13. Februar 2007.
- [16] Resnick, Paul; Zeckhauser, Richard; Friedman, Eric; et. al.: Reputation Systems: Facilitating Trust in Internet Interactions. Communications of the ACM, volume 43, issue 12, 45-48 (2000).
- [17] Sanger, Larry: Why Wikipedia Must Jettison Its Anti-Elitism. <http://www.kuro5hin.org/story/2004/12/30/142458/25>, abgefragt am 13. Februar 2007. Kuro5hin (2004).
- [18] Schoberth, Thomas; Preece, Jenny; Heinzl, Armin: Online Communities: A Longitude Analysis of Communication Activities. Working Paper in Information Systems (2003).
- [19] Tapscott, Don; Williams, Anthony D.: Wikinomics: How Mass Collaboration Changes Everything. Portfolio (2006).
- [20] Tufte, Edward R.: Beautiful Evidence. Graphics Press, 7 – 19 (2006).